

PD 11/5/01  
10/1840060  
Seg. ID 142  
Mittra R.

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2006, 16:33:29 ; Search time 15352 Seconds  
(without alignments)  
11407.939 Million cell updates/sec

Title: US-10-840-060-142

Perfect score: 3081  
Sequence: 1 aaactaaaaattgtgtgc.....gcgaatggcagcgtagt 3081

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_in.\*  
3: gb\_env.\*  
4: gb\_cm.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pr.\*  
9: gb\_ro.\*  
10: gb\_sta.\*  
11: gb\_sv.\*  
12: gb\_un.\*  
13: gb\_vi.\*  
14: gb\_hg.\*  
15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3081	100.0	3081	6	CQ573345	CQ573345 Sequence
2	3081	100.0	3081	6	AX77288	AX77288 Sequence
3	2605.2	84.6	3516	2	AY051763	AY051763 Drosophil
4	1547.8	50.2	8043	6	CQ573344	CQ573344 Sequence
5	1547.8	50.2	35708	14	AC017670	AC017670 Drosophil
6	1547.8	50.2	168583	2	AC104510	AC104510 Drosophil
7	1547.8	50.2	283075	2	AE003485	AE003485 Drosophil
8	1547.8	50.2	334028	14	AC116537	AC116537 Drosophil
9	1117.6	36.3	180919	14	AC007521	AC007521 Drosophil
10	967.8	31.4	180919	14	AC007521	AC007521 Drosophil
11	568.6	18.5	2518	2	BT001278	BT001278 Drosophil
12	564.6	18.3	2190	6	CQ577533	CQ577533 Sequence
13	564.6	18.3	4190	6	CQ577532	CQ577532 Sequence
14	564.6	18.3	84419	14	AC006245	AC006245 Drosophil
15	564.6	18.3	180361	2	AC099036	AC099036 Drosophil
16	564.6	18.3	283700	2	AE003459	AE003459 Drosophil
17	563	18.3	38036	14	AC013192	AC013192 Drosophil
18	505.8	16.4	1970	6	CQ578232	CQ578232 Sequence

19	505.8	16.4	2104	2	DQ151904	DQ151904 Drosophil
20	505.8	16.4	3970	6	CQ578231	CQ578231 Sequence
21	505.8	16.4	38036	14	AC013192	AC013192 Drosophil
22	505.8	16.4	84419	14	AC006245	AC006245 Drosophil
23	505.8	16.4	180361	2	AC099036	AC099036 Drosophil
24	505.8	16.4	181463	2	AC008352	AC008352 Drosophil
25	505.8	16.4	283700	2	AE003459	AE003459 Drosophil
26	477	15.5	2325	9	BC006841	BC006841 Mus muscu
27	464.2	15.1	3069	9	RNU44979	U44979 Rattus norv
28	461.4	15.0	2703	9	CGU11790	U11790 Cricetulus
29	454.6	14.8	1292	6	AR263768	Sequence
30	454.6	14.8	1292	6	AR263771	Sequence
31	454.6	14.8	1420	6	AR412008	Sequence
32	454.6	14.8	1421	6	AR412011	Sequence
33	454.6	14.8	2172	6	AR263774	Sequence
34	454.6	14.8	2172	6	AR412014	Sequence
35	453	14.7	2178	8	BT006759	Homo sapi
36	453	14.7	2178	11	AY889031	Synthetic
37	453	14.7	2178	11	AY891611	Synthetic
38	453	14.7	2178	11	BT007930	Synthetic
39	453	14.7	2422	8	AB072747	Macaca fa
40	453	14.7	2815	8	BC014924	Homo sapi
41	453	14.7	2823	8	BC008764	Homo sapi
42	453	14.7	2865	6	AX281799	Sequence
43	451.4	14.7	2175	8	CR450302	Homo sapi
44	449.8	14.6	1175	6	AR263769	Sequence
45	449.8	14.6	1175	6	AR412009	Sequence

ALIGNMENTS

RESULT 1  
LOCUS CQ573345 3081 bp DNA linear PAT 02-FEB-2004  
DEFINITION Sequence 1103 from Patent WO0171042.  
ACCESSION CQ573345  
VERSION CQ573345.1 GI:41637543  
KEYWORDS  
SOURCE Drosophila sp.  
ORGANISM Drosophila sp.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1  
AUTHORS Venter J.C., Adams M., Li, P.W. and Myers E.W.  
TITLE Detection kits, such as nucleic acid arrays, for detecting the  
expression of 10,000 or more Drosophila genes and uses thereof  
JOURNAL Patent: WO 0171042-A 1103 27-SEP-2001.  
PE Corporation (NY) (US)

FEATURES  
source  
1..3081  
/organism="Drosophila sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:7242"

ORIGIN

Query Match 100.0%; Score 3081; DB 6; Length 3081;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3081; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 AAACATAAAATTTGTTGCTGACATCGTTCGATTAATCGCAGATTTT 60  
Db 1 AAACATAAAATTTGTTGCTGACATCGTTCGATTAATCGCAGATTTT 60  
Qy 61 GTGATTTTCGTTGTCGTCGATTAATCGCAGATTTT 120  
Db 61 GTGATTTTCGTTGTCGTCGATTAATCGCAGATTTT 120  
Qy 121 AACACATTGTACGGTTGGAGATAGCAGAGAGAGCCAGCGGTCGCGTTCCTT 180  
Db 121 AACACATTGTACGGTTGGAGATAGCAGAGAGAGCCAGCGGTCGCGTTCCTT 180

181 TTTGTTATCCGTTGCGAGCGCAACGAAACGACAGTTGGCATTTGAATTCAGCAAAAC 240  
181 TTTGTTATCCGTTGCGAGCGCAACGAAACGACAGTTGGCATTTGAATTCAGCAAAAC 240  
241 ACACATACTAAGCCGCGACCCGCAAGCAGCAGACACACACACTGGGACACTCGAAAAA 300  
241 ACACATACTAAGCCGCGACCCGCAAGCAGCAGACACACACACTGGGACACTCGAAAAA 300  
301 AAAAAACAGACCGCTGTCGGCGACCTCGACAAGCAGTTGGTTTCGATTTAGTTGCAATGC 360  
301 AAAAAACAGACCGCTGTCGGCGACCTCGACAAGCAGTTGGTTTCGATTTAGTTGCAATGC 360  
361 CTTGAATTCGGTTCCGGGCTTAGTTTCCACAAAGTTTATCGCTCGTCAAGAAACAGGAA 420  
421 TAAAAATTATTTTCGACCTTAAAAAATCTGACTAAATTTGTTTTGTTTATGTAATTTAT 480  
421 TAAAAATTATTTTCGACCTTAAAAAATCTGACTAAATTTGTTTTGTTTATGTAATTTAT 480  
481 TAGGCACATTTTGCACACACAAACGTAGTTACTACATCTAGCACTAACGGAACCTCCT 540  
481 TAGGCACATTTTGCACACACAAACGTAGTTACTACATCTAGCACTAACGGAACCTCCT 540  
541 GCAAGCAGTGGAAAGTTGTGTCCTCATCAAGCAGTACTCGGAGTTAACGCGAGATAAGCCG 600  
601 GAGAAAGAGAAGAGATCGGTGGAGAAATAGAGATATACAGTGGAGTCAAGAGGAAGA 660  
601 GAGAAAGAGAAGAGATCGGTGGAGAAATAGAGATATACAGTGGAGTCAAGAGGAAGA 660  
661 TCATGCAATGATTAACGTGGGCGAGAGCGTCAAGATCAAGCGGACGATGCGCGCTCC 720  
661 TCATGCAATGATTAACGTGGGCGAGAGCGTCAAGATCAAGCGGACGATGCGCGCTCC 720  
721 ACATGGCCGTGGTGGCGGTGATCAACAGTCCGGCAGTGCATCACAGTGCATGGAATG 780  
721 ACATGGCCGTGGTGGCGGTGATCAACAGTCCGGCAGTGCATCACAGTGCATGGAATG 780  
781 AGCGGGCGAAACGAGGCGAAGAGGTAGAACTGGAGCGCATCTACGCTCAATCCGG 840  
781 AGCGGGCGAAACGAGGCGAAGAGGTAGAACTGGAGCGCATCTACGCTCAATCCGG 840  
841 AGCTAATCAAGATACTGTCGAAAGCAGCGCCGCGAGCCCAAGAAACAGCCACCG 900  
841 AGCTAATCAAGATACTGTCGAAAGCAGCGCCGCGAGCCCAAGAAACAGCCACCG 900  
901 CGCGGATGAACCTCTCGGCTAATCCACACAATCGGCTATCGGTGGCAATCTCACAGCC 960  
901 CGCGGATGAACCTCTCGGCTAATCCACACAATCGGCTATCGGTGGCAATCTCACAGCC 960  
961 GTATGACCATGGCGGAAACATGCTGAACAAGATCCAGAAAGCCAGTTCGATTCGCAATC 1020  
961 GTATGACCATGGCGGAAACATGCTGAACAAGATCCAGAAAGCCAGTTCGATTCGCAATC 1020  
1021 CGATTTGTCAGCAGCAATAGCTGTAATCAAAACAGCACTCCAACTACAGCGCGCGGAG 1080  
1021 CGATTTGTCAGCAGCAATAGCTGTAATCAAAACAGCACTCCAACTACAGCGCGCGGAG 1080  
1081 GTGTGGCCACCAACGTCGACGACCACTGGATTACAGCGTCCAGGTTACTCGCAAGCTG 1140  
1081 GTGTGGCCACCAACGTCGACGACCACTGGATTACAGCGTCCAGGTTACTCGCAAGCTG 1140  
1141 CTACGGGCGACAGCAGCAAGGATCGCTCGGCGTCCCTAATAACATGTCGCAATC 1200  
1141 CTACGGGCGACAGCAGCAAGGATCGCTCGGCGTCCCTAATAACATGTCGCAATC 1200  
1201 CCAGCGGCGACAGCTGCTGTCGGCGGCGACAGGAGTCCCACTCGGCGCCACACCC 1260  
1201 CCAGCGGCGACAGCTGCTGTCGGCGGCGACAGGAGTCCCACTCGGCGCCACACCC 1260  
1261 AGGAGCTGGCGCGCTAGTAGTACCGGGGATCGCACGCAATGAAAGAGGTGGAGCGACTGA 1320

1261 AGGAGCTGGCGCGCTAGTAGTACCGGGGATCGCACGCAATGAAAGAGGTGGAGCGACTGA 1320  
1321 AGGAGATTCGGAGAGCGACGCGCCGACAGCGCGAGATGAAAGAGAGAGGTGGCGC 1380  
1321 AGGAGATTCGGAGAGCGACGCGCCGACAGCGCGAGATGAAAGAGAGAGGTGGCGC 1380  
1381 TGATGAACAGAGATCCGGGCAATCAAACTGGGAGCGGCAATGATACGCAATATC 1440  
1381 TGATGAACAGAGATCCGGGCAATCAAACTGGGAGCGGCAATGATACGCAATATC 1440  
1441 AGAGCAGCTGGAAATTTGTGCGCTGCTCGATGGCAGGCCGTGATGACCATCAGATCA 1500  
1441 AGAGCAGCTGGAAATTTGTGCGCTGCTCGATGGCAGGCCGTGATGACCATCAGATCA 1500  
1501 CAGTGTGCTGGCAAGCGTCCATTTAGCCGAAGGAGTCAATCGCAAGGAGATCGATG 1560  
1501 CAGTGTGCTGGCAAGCGTCCATTTAGCCGAAGGAGTCAATCGCAAGGAGATCGATG 1560  
1561 TCATTTGCTGGCGCAAGGACATGCTCATCGTCGACGCGCGCAGCAAGGTGCGACC 1620  
1561 TCATTTGCTGGCGCAAGGACATGCTCATCGTCGACGCGCGCAGCAAGGTGCGACC 1620  
1621 TCACCAAGTTCTTGGAGAACCAAGTTTTCGCTTCGACTAGCCCTTCAACGACAGTGG 1680  
1621 TCACCAAGTTCTTGGAGAACCAAGTTTTCGCTTCGACTAGCCCTTCAACGACAGTGG 1680  
1681 ACAATGCCATGGTATACAAATACACAGCAAGCGTTGGTGAACCAATTTTCAGAGGCG 1740  
1681 ACAATGCCATGGTATACAAATACACAGCAAGCGTTGGTGAACCAATTTTCAGAGGCG 1740  
1741 GAATGGGACGCTGCTCGCTTACCGCCAGCGGATCGGGCAACGACACACCTGCGCG 1800  
1741 GAATGGGACGCTGCTCGCTTACCGCCAGCGGATCGGGCAACGACACACCTGCGCG 1800  
1801 GTGAGTTTAATFGAAAGTGCAGGACTGCAAGAACGCGCATCTACGCCATGCGGCAAG 1860  
1801 GTGAGTTTAATFGAAAGTGCAGGACTGCAAGAACGCGCATCTACGCCATGCGGCAAG 1860  
1861 ATGCTTTTGTGACCTGAAATAGCGGTTTACCGCGCATGAATCTAGTCTGCGGCA 1920  
1861 ATGCTTTTGTGACCTGAAATAGCGGTTTACCGCGCATGAATCTAGTCTGCGGCA 1920  
1921 GTTTCTTTGAGATTTACAGTGGCAAGTCTTCGATCTTCGTCGACAGCAAGCAAACTGC 1980  
1921 GTTTCTTTGAGATTTACAGTGGCAAGTCTTCGATCTTCGTCGACAGCAAGCAAACTGC 1980  
1981 GCGTCTTGGAGATGGTAAACAGCAAGTGCAGGTGGTGGGACTCACCGAGAGGTGTCG 2040  
1981 GCGTCTTGGAGATGGTAAACAGCAAGTGCAGGTGGTGGGACTCACCGAGAGGTGTCG 2040  
2041 ATGGCGTCGAGGAGTACTGAGCTCATCCAGCAGCGCAATGCTGCCGGAATCCCGCC 2100  
2041 ATGGCGTCGAGGAGTACTGAGCTCATCCAGCAGCGCAATGCTGCCGGAATCCCGCC 2100  
2101 AGAGCTCGGCAACTCAATTTGTCGCTGCGACGCGCTTTTCCAGATTCGTCGCGCG 2160  
2101 AGAGCTCGGCAACTCAATTTGTCGCTGCGACGCGCTTTTCCAGATTCGTCGCGCG 2160  
2161 CGCAGGCTCGACGAAAGATCCATGGCAAGTTCTCGTTTCATCGATCTGCGGCGCAATGAGC 2220  
2161 CGCAGGCTCGACGAAAGATCCATGGCAAGTTCTCGTTTCATCGATCTGCGGCGCAATGAGC 2220  
2221 GGGCGTGGGACACTTCTCGGCGCATCGGAGACGCGTATGGAGGGTGCAGATTAAACA 2280  
2221 GGGCGTGGGACACTTCTCGGCGCATCGGAGACGCGTATGGAGGGTGCAGATTAAACA 2280  
2281 AATCGCTGCTGGCGCTCAAGGAGTGCATTCGTGGTTGGGCAACAGTTCGCGCCACTTGC 2340  
2281 AATCGCTGCTGGCGCTCAAGGAGTGCATTCGTGGTTGGGCAACAGTTCGCGCCACTTGC 2340  
2341 CTTTCGCTGCTCAAACTCAACGAGTGTGCGGATCTGTTTCATTGGCGGAGAGCA 2400



QY	1021	CGATTGTCAGCAGCAATAGCGTGAATACAAACAGCAACTCCAACTACGCGCGCGGAG	1080
DB			
QY	1021	CGATTGTCAGCAGCAATAGCGTGAATACAAACAGCAACTCCAACTACGCGCGCGGAG	1080
DB			
QY	1081	GTGTTGGCACCAACAGTCGACGACCACTGGATTACAGCGTCCACGGTACTCGCAAGCTG	1140
DB			
QY	1081	GTGTTGGCACCAACAGTCGACGACCACTGGATTACAGCGTCCACGGTACTCGCAAGCTG	1140
DB			
QY	1141	CTACCGGCGCAGCAGCAGCAAGGATCGCTCGCGGTGCCTTAATACATATTGCCCAATC	1200
DB			
QY	1141	CTACCGGCGCAGCAGCAGCAAGGATCGCTCGCGGTGCCTTAATACATATTGCCCAATC	1200
DB			
QY	1201	CCAGCGGCGCAGCAGCTGCTGGTCGCGGCGCACAGGAGTCGCCACTCGCGGCGACACCC	1260
DB			
QY	1201	CCAGCGGCGCAGCAGCTGCTGGTCGCGGCGCACAGGAGTCGCCACTCGCGGCGACACCC	1260
DB			
QY	1261	AGGGAGCTGGCGGCGTAGTACCGCGGATCGCAGCGATTGAAGAGGTGGAGCGACTGA	1320
DB			
QY	1261	AGGGAGCTGGCGGCGTAGTACCGCGGATCGCAGCGATTGAAGAGGTGGAGCGACTGA	1320
DB			
QY	1321	AGGAGAATCGCGAGAGGAGCGCGCCCGACAGGCCGAGATGAAGAGGAGAGGTGGCGC	1380
DB			
QY	1321	AGGAGAATCGCGAGAGGAGCGCGCCCGACAGGCCGAGATGAAGAGGAGAGGTGGCGC	1380
DB			
QY	1381	TGATGAACCAAGATCCCGGCAATCCAACTCGGGAGCGGCCAAATGATACGCGAATATC	1440
DB			
QY	1381	TGATGAACCAAGATCCCGGCAATCCAACTCGGGAGCGGCCAAATGATACGCGAATATC	1440
DB			
QY	1441	AGACGACGCTGGAATTTGTGCGCGTGTGATGTGCCAGGCGCTCGATGACCATCAGATCA	1500
DB			
QY	1441	AGACGACGCTGGAATTTGTGCGCGTGTGATGTGCCAGGCGCTCGATGACCATCAGATCA	1500
DB			
QY	1501	CAGTGTGCGTGGCGCAAGCTGCCATTAGCGCAAGGAGTCAATCGCAAGGAGATCGATG	1560
DB			
QY	1501	CAGTGTGCGTGGCGCAAGCTGCCATTAGCGCAAGGAGTCAATCGCAAGGAGATCGATG	1560
DB			
QY	1561	TCATTTTCGTTGCGCGCAAGGACATGCTCATCTGTGCAGGACCGCGCAGCAAGTGCACC	1620
DB			
QY	1561	TCATTTTCGTTGCGCGCAAGGACATGCTCATCTGTGCAGGACCGCGCAGCAAGTGCACC	1620
DB			
QY	1621	TCACCAAGTTCTTCGGAGAACCAAAATTCGCTTCGACTACGCTTCAACGACACAGTGGC	1680
DB			
QY	1621	TCACCAAGTTCTTCGGAGAACCAAAATTCGCTTCGACTACGCTTCAACGACACAGTGGC	1680
DB			
QY	1681	ACAATGCCATGGTTATACAAATACACAGCCAAAGCGTTGGTGAACCAATTTTCGAGGGC	1740
DB			
QY	1681	ACAATGCCATGGTTATACAAATACACAGCCAAAGCGTTGGTGAACCAATTTTCGAGGGC	1740
DB			
QY	1741	GAATGGCGAGTGTCTTCGCTTCGCGCAGACGGGATCGGCAAAACGCAACCATGGCGC	1800
DB			
QY	1741	GAATGGCGAGTGTCTTCGCTTCGCGCAGACGGGATCGGCAAAACGCAACCATGGCGC	1800
DB			
QY	1801	GTGAGTTTAATGAAAGGTCAGGACTGCAAGAACCGGCATCTACGCCCATGGCGGCCAAGG	1860
DB			
QY	1801	GTGAGTTTAATGAAAGGTCAGGACTGCAAGAACCGGCATCTACGCCCATGGCGGCCAAGG	1860
DB			
QY	1861	ATGTCTTTGTGACCCCTGAATATACCGCGTTACCGCGCATGAATCTAGTGTCTCGGCCA	1920
DB			
QY	1861	ATGTCTTTGTGACCCCTGAATATACCGCGTTACCGCGCATGAATCTAGTGTCTCGGCCA	1920
DB			
QY	1921	GTTTCTTTTGAGATTTACAGTGGCAAGTCTTCGATCTTCTGTCGCAAGACGAGAACTGC	1980
DB			
QY	1921	GTTTCTTTTGAGATTTACAGTGGCAAGTCTTCGATCTTCTGTCGCAAGACGAGAACTGC	1980
DB			
QY	1981	CGCTCTCTGGAGGATGGTAAACAGCAAGTCAGGTGGTGGGACTCACCGAGAGGTCGTTCG	2040
DB			
QY	1981	CGCTCTCTGGAGGATGGTAAACAGCAAGTCAGGTGGTGGGACTCACCGAGAGGTCGTTCG	2040
DB			
QY	2041	ATGCGTCTGAGGAGTACTGAAGCTATCAGCAAGGCAATGTGCGCGCAATCTCGGCC	2100
DB			
QY	2041	ATGCGTCTGAGGAGTACTGAAGCTATCAGCAAGGCAATGTGCGCGCAATCTCGGCC	2100
DB			

Qy	2101	AGAGCTCGGCCAACTCCAAATTGTCGCGTTCGCAACGCCGTTTTCAGATATGTGCTCGCGC	2160
Db	2101	AGACGTGCGGCCAACTCCAAATTGTCGCGTTCGCAACGCCGTTTTCAGATATGTGCTCGCGC	2160
Qy	2161	CGCAGGCTCGACGAAGATCCATGGCAAGTTCTCGTTTCATCGATCTGGCGGGCAATGAGC	2220
Db	2161	CGCAGGCTCGACGAAGATCCATGGCAAGTTCTCGTTTCATCGATCTGGCGGGCAATGAGC	2220
Qy	2221	GGGCGGTGGAACATTCTCGGCCGATCGGCAGACGGGTATGGAGGTTGCCGAGATTAACA	2280
Db	2221	GGGCGGTGGAACATTCTCGGCCGATCGGCAGACGGGTATGGAGGTTGCCGAGATTAACA	2280
Qy	2281	AATCGCTGCTCGGCCCTCAAGAGTGCAATTCGTGCTTGGCAACAAGTCGGGCCACTTGC	2340
Db	2281	AATCGCTGCTCGGCCCTCAAGAGTGCAATTCGTGCTTGGCAACAAGTCGGGCCACTTGC	2340
Qy	2341	CTTTCGTGTCTCCAAACTCACCAGAGTCTGCGCGACTCGTTCAATTGGCGAAGACGA	2400
Db	2341	CTTTCGTGTCTCCAAACTCACCAGAGTCTGCGCGACTCGTTCAATTGGCGAAGACGA	2400
Qy	2401	AGAGTCGATGATAGCAATGATCTCGCGGGCACTTATAGTCTCTCGAGACAACGCTCAACA	2460
Db	2401	AGAGTCGATGATAGCAATGATCTCGCGGGCACTTATAGTCTCTCGAGACAACGCTCAACA	2460
Qy	2461	CGCTGCGCTATCGGATCGTGTCACAGAGCTGCTGCTCAAGGATATCTCGAAGTTTGC	2520
Db	2461	CGCTGCGCTATCGGATCGTGTCACAGAGCTGCTGCTCAAGGATATCTCGAAGTTTGC	2520
Qy	2521	CTGGCGGACACCGGAGCCATCGAGATCACGGACGACGAGGAGGAGGAGCTCAACA	2580
Db	2521	CTGGCGGACACCGGAGCCATCGAGATCACGGACGACGAGGAGGAGGAGCTCAACA	2580
Qy	2581	TGTTGCATCCGCACTCGCATCAGCTGCATCCCAATTCGCAATGCAACGGCGAGCGAGTCSA	2640
Db	2581	TGTTGCATCCGCACTCGCATCAGCTGCATCCCAATTCGCAATGCAACGGCGAGCGAGTCSA	2640
Qy	2641	ATAATCAGCGTGTCTCGGCTCTCATCTCGGGGCGTTCATTCCGACACAATAATAATA	2700
Db	2641	ATAATCAGCGTGTCTCGGCTCTCATCTCGGGGCGTTCATTCCGACACAATAATAATA	2700
Qy	2701	ACAACAACGAACCGGAAACGCGGCAACATGCACTTGCCCATGTGAGTTGCTGAGCG	2760
Db	2701	ACAACAACGAACCGGAAACGCGGCAACATGCACTTGCCCATGTGAGTTGCTGAGCG	2760
Qy	2761	AACACGAGTTCGACGAGCTGATTTGTCAGACAACAGGCCATCGACGACCTGCGAGCA	2820
Db	2761	AACACGAGTTCGACGAGCTGATTTGTCAGACAACAGGCCATCGACGACCTGCGAGCA	2820
Qy	2821	CGGAGGAGATGGTGGTGAGTATCATCGCACCGTTAATGCCACACTGGAGACCTTCTCG	2880
Db	2821	CGGAGGAGATGGTGGTGAGTATCATCGCACCGTTAATGCCACACTGGAGACCTTCTCG	2880
Qy	2881	CCGAGTTCGAAGCGCTGTACAATCTGACCAACTATGTGGACTACGACGAGGACTCGTACT	2940
Db	2881	CCGAGTTCGAAGCGCTGTACAATCTGACCAACTATGTGGACTACGACGAGGACTCGTACT	2940
Qy	2941	GCAACCGGCGGAGTCGATGTTCTCGACGCTCGTGACATCGCCATCCAGTCCGCGACA	3000
Db	2941	GCAACCGGCGGAGTCGATGTTCTCGACGCTCGTGACATCGCCATCCAGTCCGCGACA	3000
Qy	3001	TGATGGCCGAATATCGCGCAAGTTGGCCAAGGAGGAGATGCTGTGTCGACGCTTCAATT	3060
Db	3001	TGATGGCCGAATATCGCGCAAGTTGGCCAAGGAGGAGATGCTGTGTCGACGCTTCAATT	3060
Qy	3061	CGCCGAATGGCAAGCGTTAGT	3081
Db	3061	CGCCGAATGGCAAGCGTTAGT	3081

RESULT 3  
AY051763  
LOCUS  
DEFINITION

AY051763 *Drosophila melanogaster* LD29208 full length cDNA. 3516 bp mRNA linear INV 27-AUG-2001



ACCESSION AY051763  
VERSION AY051763.1  
KEYWORDS GI:15291836  
SOURCE FLI\_CDNA  
ORGANISM Drosophila melanogaster (fruit fly)

REFERENCE  
AUTHORS Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
1. (bases 1 to 3516)  
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,  
Champe, M., Chavez, C., Dorsett, V., Farfan, D., Frise, B., George, R.,  
Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J.,  
Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K.,  
Yu, C., Lewis, S. E., Rubin, G. M. and Celniker, S.  
Direct Submission  
Submitted (10-AUG-2001) Berkeley Drosophila Genome Project,  
Lawrence Berkeley National Laboratory, One Cyclotron Road,  
Berkeley, CA 94720, USA  
Sequence submitted by:  
Berkeley Drosophila Genome Project  
Lawrence Berkeley National Laboratory  
Berkeley, CA 94720  
This clone was sequenced as part of a high-throughput process to  
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,  
Science 2000). The sequence has been subjected to integrity checks  
for sequence accuracy, presence of a polyA tail and contiguity  
within 100 kb in the genome. Thus we believe the sequence to  
reflect accurately this particular cDNA clone. However, there are  
artifacts associated with the generation of cDNA clones that may  
have not been detected in our initial analyses such as internal  
priming, priming from contaminating genomic DNA, retained introns  
due to reverse transcription of unspliced precursor RNAs, and  
reverse transcriptase errors that result in single base changes.  
For further information about this sequence, including its location  
and relationship to other sequences, please visit our Web site  
(<http://fruitfly.berkeley.edu>) or send email to  
[cdna@fruitfly.berkeley.edu](mailto:cdna@fruitfly.berkeley.edu).

FEATURES  
source location/Qualifiers  
1..3516  
/organism="Drosophila melanogaster"  
/mol\_type="mRNA"  
/strain="y; cn bw sp"  
/db\_xref="taxon:7227"  
/map="10A6-10A7"  
1..3516  
/genes="Klp10A"  
/notes="alignment with genomic scaffold AE003485"  
/db\_xref="FLYBASE:FBgn0030268"  
337..2754  
/genes="Klp10A"  
/notes="Longest ORF"  
/codon\_start=1  
/product="Lb29208p"  
/protein\_id="AAK93187.1"  
/db\_xref="GI:15291837"  
/db\_xref="FLYBASE:FBgn0030268"  
/translation="MDMITVGSQSVKIKRTDGRVHMAVAVINOSGKCIIVWYERGET  
KGEVELDAILTNPELMQDTVEHAAPEPKQATAPMNLNRNPTQSAIGNLTSMRT  
MAGNMLTKIQESQIPNPIVSNSVNTNSNTAGCGGGTTTSTTGLRPRYSOAA  
TQCOOTRIASAVPNTLPNSAAAGPAAQGVATAATTQAGAGASTRESHAKVER  
LAKNEKRARQAEKKEKVALMNQDPGNPWETAQIMREYSTLEFVPLLDQAVDD  
HQITCVKRPISRKENVREIDVISVPRKDWLI VHEPRSKVDLTFLFNHFRFDYA  
FNDCDNAMPYKYTKAPLVKTI PEGGMATCFPAYGTGSGKTHMGEEFNKQVDCNG  
IYAMAKQVFDVFNPRYAMNLVWSASF EYISGKVFLLSDKQRLVLEDKQVQ  
VVLGTEKVDVGEVVKLIQHGNAASTSGTSNSRSASHVFOI VLRPOGSTKIHG  
KFSFDLAKNERGVDTSSADROTMEAGIEINKSLALKEICIRALKQSHLPRVSKL  
TQVLRDSFGEKSKMIAMI SPGLSSCEHTLNTLYADRVKELVVKDI VECVPGDIT  
SPIETDDEESELNVPVHSHQHPNSHAPASQNNORAPASHSHGVVHNNNNN  
KNGNAGMLDLMLSSISEHMSDELIVHOAIDDLQOETEMVVEYHRTVNALETFLA  
ESKALYNLTNYVDYDQDSYCKRGESMFSQLDIAIQCRDMMAEYRAKLEEMLSCSF  
NSPNGR"

ORIGIN

Query Match 84.6%; Score 2605.2; DB 2; Length 3516;  
Best Local Similarity 98.8%; Pred. No. 0;  
Matches 2625; Conservative 0; Mismatches 33; Indels 0; Gaps 0;  
QY 424 AATTATTTTCGACCTAAATAATCTGACTAAATTTGTGTTTTTGTATGTTATTTAG 483  
DB 98 AGTTTTTTTTTAAATGGAAAGTCTAAACATTTGTAGCGTTGGAAAGATAGCGGAAG 157  
QY 484 GCACATTTTGCACACCAACAGTAGTTACTACATCTACGACTAAACGGAACTCTCTCTGCA 543  
DB 158 GCACATTTTGCACACCAACAGTAGTTACTACATCTACGACTAAACGGAACTCTCTCTGCA 217  
QY 544 AGCAGTGGAAAGTTGCTGTCATCAAGCAGTACTCGGAGTTAAAGCAGGATAAGCCGGAG 603  
DB 218 AGCAGTGGAAAGTTGCTGTCATCAAGCAGTACTCGGAGTTAAAGCAGGATAAGCCGGAG 277  
QY 604 AAAGACAAGAGAGATCGTGGAGAAATAGAGATATACAGTGTGAGTCAAAAGAGGAGGATCA 663  
DB 278 AAAGAGAAGAGATCGTGGAGAAATAGAGATATACAGTGTGAGTCAAAAGAGGAGGATCA 337  
QY 664 TGGACATGATTACGTTGGGGCAGAGCGTCAAGATCAAGCGGACGGATGGCCGCTCCACA 723  
DB 338 TGGACATGATTACGTTGGGGCAGAGCGTCAAGATCAAGCGGACGGATGGCCGCTCCACA 397  
QY 724 TGGCCGTGTTGGCGGTGATCAACCACTCGGGCAAGTGCATCACAGTCGAATGTGACGAGC 783  
DB 398 TGGCCGTGTTGGCGGTGATCAACCACTCGGGCAAGTGCATCACAGTCGAATGTGACGAGC 457  
QY 784 GCGGCCAAAAGAGGCAAGGAGGTAGAACTGAGACGCCATCTACTCAGCTCAATCCGAGC 843  
DB 458 GCGGCCAAAAGAGGCAAGGAGGTAGAACTGAGACGCCATCTACTCAGCTCAATCCGAGC 517  
QY 844 TAATGCAAGATACTGTGCAACAGCACGCCGCCCGGAGCCCAAGAAACCAAGCACCAGCGC 903  
DB 518 TAATGCAAGATACTGTGCAACAGCACGCCGCCCGGAGCCCAAGAAACCAAGCACCAGCGC 577  
QY 904 CGATGAACCTCTCGCGTAAATCCACAACTCGGCTATCGGTGGCAATCTCACAGCCGTA 963  
DB 578 CGATGAACCTCTCGCGTAAATCCACAACTCGGCTATCGGTGGCAATCTCACAGCCGTA 637  
QY 964 TGACCATGCGCGGAAACATGCTGAACAAGATCAGAAAGCCAGTGCATTTCCCAATCCGA 1023  
DB 638 TGACCATGCGCGGAAACATGCTGAACAAGATCAGAAAGCCAGTGCATTTCCCAATCCGA 697  
QY 1024 TTGTGACAGCAATAGCGTGAATACAAACAGCAACTCCAACTACGCGCGCGGAGGTG 1083  
DB 698 TTGTGACAGCAATAGCGTGAATACAAACAGCAACTCCAACTACGCGCGCGGAGGTG 757  
QY 1084 GTGGCACCAACAGCTCGACGACCACTGGATTACAGGTTACAGGTTACTCGCAAGCTGCTA 1143  
DB 758 GTGGCACCAACAGCTCGACGACCACTGGATTACAGGTTACAGGTTACTCGCAAGCTGCTA 817  
QY 1144 CCGGCCAGCAGCAGACAAGGATCGCTCGCGCGTGCCTTAATACACATTTGCCCAATCCCA 1203  
DB 818 CCGGCCAGCAGCAGACAAGGATCGCTCGCGCGTGCCTTAATACACATTTGCCCAATCCCA 877  
QY 1204 GCGCGCAGCAGTGTGTTGTCGCGGCGCAAGAGAGTGCCTACCTGCGGCGCACAAACCCAGG 1263  
DB 878 GCGCGCAGCAGTGTGTTGTCGCGGCGCAAGAGAGTGCCTACCTGCGGCGCACAAACCCAGG 937  
QY 1264 GAGCTGGCGCGCTAGTATCCCGCGGATCGCAAGATTGAAGAGGTGGAGCGCTGAAGG 1323  
DB 938 GAGCTGGCGCGCTAGTATCCCGCGGATCGCAAGATTGAAGAGGTGGAGCGCTGAAGG 997  
QY 1324 AGAATCGGAGAGGAGCGCGCGCGGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1383  
DB 998 AGAATCGGAGAGGAGCGCGCGCGGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1057  
QY 1384 TGAACCAAGGATCCGGGCAATCAAACTGGGAGACGCGCAAAATGATACGCAATATCAGA 1443  
DB 1058 TGAACCAAGGATCCGGGCAATCAAACTGGGAGACGCGCAAAATGATACGCAATATCAGA 1117  
QY 1444 GCACGCTGGAAATTTGTGCGCTGCTCGATGGCCAGCGCGCTCGATGACATGACATGACAG 1503

Db	1118	GCACGCTGGAAATTTGTGCGCGTGTCTCGATGGCCAGGCGGTGAGTACCATCATGATCAACAG	1177
Qy	1504	TGTGCGTGCACGAAGCGTCCCAATTAGCCGCAAGAGAGGTCAATTCGCAAGGAGATCGATGTCA	1563
Db	1178	TGTGCGTGCACGAAGCGTCCCAATTAGCCGCAAGAGAGGTCAATTCGCAAGGAGATCGATGTCA	1237
Qy	1564	TTTTCGTTGCGCGCAAGAGCATGCTCTCATTCGTGTCACGAGCGCGCAGCAAGGTCGACCTCA	1623
Db	1238	TTTTCGTTGCGCGCAAGAGCATGCTCTCATTCGTGTCACGAGCGCGCAGCAAGGTCGACCTCA	1623
Qy	1624	CCAAAGTTCCCTGGAGAAACCAAGTTTTTCGCTTCGACTAGCGCTTCAACGACACGCTGGAC	1683
Db	1298	CCAAAGTTCCCTGGAGAAACCAAGTTTTTCGCTTCGACTAGCGCTTCAACGACACGCTGGAC	1357
Qy	1684	ATGCCATGGTATACAAATACACAGCCAAAGCGTTTGGTGAATAACATTTTCGAGGGCGGAA	1743
Db	1358	ATGCCATGGTATACAAATACACAGCCAAAGCGTTTGGTGAATAACATTTTCGAGGGCGGAA	1417
Qy	1744	TGGCGACGTGCTTCGCTTACCGCCAGACGGGATCGGGCAAAACGCAACCATTCGAGGGCGG	1803
Db	1418	TGGCGACGTGCTTCGCTTACCGCCAGACGGGATCGGGCAAAACGCAACCATTCGAGGGCGG	1477
Qy	1804	AGTTTAATGGAAAGGTGACGACCTGCAAGAAACGGCATCTAGCCCATGGCGGCAAGGATG	1863
Db	1478	AGTTTAATGGAAAGGTGACGACCTGCAAGAAACGGCATCTAGCCCATGGCGGCAAGGATG	1537
Qy	1864	TCCTTTGTGACCCCTCAATATGCGCGTTTACCGCGCATGAATCTTAGTCTGCTCGGCCAGTT	1923
Db	1538	TCCTTTGTGACCCCTCAATATGCGCGTTTACCGCGCATGAATCTTAGTCTGCTCGGCCAGTT	1597
Qy	1924	TCCTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTCTGTCGCAAGCAAGAACTCGCG	1983
Db	1598	TCCTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTCTGTCGCAAGCAAGAACTCGCG	1657
Qy	1984	TCCTTGGAGGATGTTAAACAGCAAGTGCAGTGTGGGACTCAACGAGAAAGTGGTCAATG	2043
Db	1658	TCCTTGGAGGATGTTAAACAGCAAGTGCAGTGTGGGACTCAACGAGAAAGTGGTCAATG	1717
Qy	2044	CGCTCGAGGAGGTACTCAAGCTCATCAGCACCGCATGCTGCCGCAATCGGCCAGCA	2103
Db	1718	CGCTCGAGGAGGTACTCAAGCTCATCAGCACCGCATGCTGCCGCAATCGGCCAGCA	1777
Qy	2104	CGTCGGCCAACTCCAAATTCGTTCGCGTTTCGCACGCGGTTTTCAGATTTGCTCGCGCCG	2163
Db	1778	CGTCGGCCAACTCCAAATTCGTTCGCGTTTCGCACGCGGTTTTCAGATTTGCTCGCGCCG	1837
Qy	2164	AGGCTCGACGAGATCCATGGCAAGTTCTCGTTTCATTCGATCTGGCGGCAATGAGCGG	2223
Db	1838	AGGCTCGACGAGATCCATGGCAAGTTCTCGTTTCATTCGATCTGGCGGCAATGAGCGG	1897
Qy	2224	CGCTGGACACTTCTCGCGCGATCGGCAGACGGGTATGGAGGTGCGGAGATTAACAAT	2283
Db	1898	CGCTGGACACTTCTCGCGCGATCGGCAGACGGGTATGGAGGTGCGGAGATTAACAAT	1957
Qy	2284	CGCTGCTGGCCCTCAAGAGTGCATTCGTTCGCTTGGCAAAACAGTCGGCCACTTGCCT	2343
Db	1958	CGCTGCTGGCCCTCAAGAGTGCATTCGTTCGCTTGGCAAAACAGTCGGCCACTTGCCT	2017
Qy	2344	TCCGTGTCTCAAACTCACCCAGGTGCTGCGACCTGTTTCATTTGGCGAGAGAGCAAGA	2403
Db	2018	TCCGTGTCTCAAACTCACCCAGGTGCTGCGACCTGTTTCATTTGGCGAGAGAGCAAGA	2077
Qy	2404	CGTGCATGATAGCCATGATCTCCCGGGACTTAGCTTCTCGGAGCACACGCTCAACACG	2463
Db	2078	CGTGCATGATAGCCATGATCTCCCGGGACTTAGCTTCTCGGAGCACACGCTCAACACG	2137
Qy	2464	TGCGCTATGCGGATCTGTGTCGAGGACTGGTGGTCAAGGATATCGTCGAAGTTTGGCCTG	2523
Db	2138	TGCGCTATGCGGATCTGTGTCGAGGACTGGTGGTCAAGGATATCGTCGAAGTTTGGCCTG	2197
Qy	2524	CGCGCGACACCGAGCCCATCGATCAACGACGAGGAGGAGGAGCTCAACATGG	2583

D	b	2198	GCGCGCAGACCGAGCCCATCGAGATTCACGACGACGAGGAGGAGGAGCTCAACATGG	2255
Q	y	2584	TGCATCCGCACTCGCATCAGCTGCATCCCAATTTCGATGCACCGCGCAGCGACTCGAATA	2643
D	b	2258	TGCATCCGCACTCGCATCAGCTGCATCCCAATTTCGATGCACCGCGCAGCGACTCGAATA	2317
Q	y	2644	ATCAGCGTGTCTCGGCCCTCTCATCACCTCGGGGCGGTCAITTCACAACAATAATAACA	2703
D	b	2318	ATCAGCGTGTCTCGGCCCTCTCATCACCTCGGGGCGGTCAITTCACAACAATAATAACA	2377
Q	y	2704	ACAACAAGACGGAAACCGCGGGAACATGAGCACTCGGCATGCTGAGTTGCTGAGCGAAC	2763
D	b	2378	ACAACAAGACGGAAACCGCGGGAACATGAGCACTCGGCATGCTGAGTTGCTGAGCGAAC	2437
Q	y	2764	ACGAGATGTCCGACGAGCTGATGTGCAGCACACAGGCCCATCGACGACCTGCAGCAGCG	2823
D	b	2438	ACGAGATGTCCGACGAGCTGATGTGCAGCACACAGGCCCATCGACGACCTGCAGCAGCG	2497
Q	y	2824	AGGAGATGTGTGGAGTATCATCGCACACCCTTAATGCCACACTGGAGACCTTCTCGCCG	2883
D	b	2498	AGGAGATGTGTGGAGTATCATCGCACACCCTTAATGCCACACTGGAGACCTTCTCGCCG	2557
Q	y	2884	AGTCGAAGGCGCTGTACAATCTGACCAAATPATGTGGACTACGACAGGACTCTGATGCA	2943
D	b	2558	AGTCGAAGGCGCTGTACAATCTGACCAAATPATGTGGACTACGACAGGACTCTGATGCA	2617
Q	y	2944	AACGGGCGAGTCGATGTCTCCACGCTGTCGACATCGCCATCCAGTCCAGTCCGCGCATGA	3003
D	b	2618	AACGGGCGAGTCGATGTCTCCACGCTGTCGACATCGCCATCCAGTCCAGTCCGCGCATGA	2677
Q	y	3004	TGGCCGAATATCGCGCAAGTTTGCCCAAGGAGAGATGCTGTGTCGAGCTTCAATTCGC	3063
D	b	2678	TGGCCGAATATCGCGCAAGTTTGCCCAAGGAGAGATGCTGTGTCGAGCTTCAATTCGC	2737
Q	y	3064	CGAATGCAAGCGTTAGT	3081
D	b	2738	CGAATGCAAGCGTTAGT	2755
RESULT 4				
COSY73344				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
Drosophila sp.				
Drosophila sp.				
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
Neoptera; Endopterygota; Diptera; Brachycera, Muscomorpha;				
Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE				
Venter, J.C., Adams, M., Li, P.W. and Myers, E.W.				
Detection kits, such as nucleic acid arrays, for detecting the				
expression of 10,000 or more Drosophila genes and uses thereof				
Patent: WO 0171042-A 1102 27-SEP-2001;				
PE Corporation (NY) (US)				
FEATURES				
source				
1..8043				
/organism="Drosophila sp."				
/mol_type="unassigned DNA"				
/db_xref="taxon:7242"				
ORIGIN				
Query Match				
Best Local Similarity				
Matches 1552; Conservative				
Score 1547.8; DB 6; Length 8043;				
Pred. No. 2.4e-286;				
Mismatches 0; Indels 0; Gaps 0;				
Q	y	156	AGCCAGGGGCTCCGTTTTTCTTTTTTTTATTCGTTGCCAGACGCAACGAAAACGACA	215
D	b	1599	AGCCAGGGGCTCCGTTTTTCTTTTTTTTATTCGTTGCCAGACGCAACGAAAACGACA	1658
Q	y	216	GTGTGGCATTTGAATTCAGCACAAACACATACTACGCGCGACCCCGCAAGCAGCACAC	275

Db 1659 GTTGGCATTTGAAATTCAGCAAAACACACATACACCGCCGACCCGCAAGCAGCACAC 1718  
QY 276 ACACACACTGGGACACTCGAAAAAACAACAGACGCTGTCGGCGACCTCGCAACAGCAG 335  
Db 1719 ACACACACTGGGACACTCGAAAAAACAACAGACGCTGTCGGCGACCTCGCAACAGCAG 1778  
QY 336 TTCGGTTCGAAATTTAGTTGTCATGCTTGAATTCGGTTCCGGTTCGGGCTTAGTTTCACAAGTT 395  
Db 1779 TTCGGTTCGAAATTTAGTTGTCATGCTTGAATTCGGTTCCGGTTCGGGCTTAGTTTCACAAGTT 1838  
QY 396 TATCGCTCGTCAAGAAACAACAGAAATAAAATTTATTTTCGACCTAAATAATCTCACTAAAT 455  
Db 1839 TATCGCTCGTCAAGAAACAACAGAAATAAAATTTATTTTCGACCTAAATAATCTCACTAAAT 1898  
QY 456 TGTGTTTTTTTATGATTTATTTATTTAGGCACATTTTGCACACACCAACGTAAGTACTAC 515  
Db 1899 TGTGTTTTTTTATGATTTATTTATTTAGGCACATTTTGCACACACCAACGTAAGTACTAC 1958  
QY 516 ATCTAGACTAACCGGAACCTCTCTCGAAGCAGTGAAGTTGCTGTCATCAAGCAAGTAC 575  
Db 1959 ATCTAGACTAACCGGAACCTCTCTCGAAGCAGTGAAGTTGCTGTCATCAAGCAAGTAC 2018  
QY 576 TCGGAGTTAACGCAAGATAAGCCGCGAGAAAGAGAGATCGGTGGAGATAGAGATA 635  
Db 2019 TCGGAGTTAACGCAAGATAAGCCGCGAGAAAGAGAGATCGGTGGAGATAGAGATA 2078  
QY 636 TACAGTGGAGTCAAGAGAGAGAGATCATGGACATGATTACGGTGGGGCAGAGCGTCAAG 695  
Db 2079 TACAGTGGAGTCAAGAGAGAGAGATCATGGACATGATTACGGTGGGGCAGAGCGTCAAG 2138  
QY 696 ATCAAGCGGACGATGGCGGCTGCACATGCGCGTGGTGGCGGTGATCAACAGTCCGGC 755  
Db 2139 ATCAAGCGGACGATGGCGGCTGCACATGCGCGTGGTGGCGGTGATCAACAGTCCGGC 2198  
QY 756 AAGTGCATCACAGTCAAGTGTACGAGCGCGGCAACGAGGCGAGGAGGTAGAAC 815  
Db 2199 AAGTGCATCACAGTCAAGTGTACGAGCGCGGCAACGAGGCGAGGAGGTAGAAC 2258  
QY 816 GAGCGCATACTCAGCTCAATCCGAGAGTAAATGCAAGATACTGTGCAACGAGCAGCGCGC 875  
Db 2259 GAGCGCATACTCAGCTCAATCCGAGAGTAAATGCAAGATACTGTGCAACGAGCAGCGCGC 2318  
QY 876 CCGGAGCCCAAGAAACAAGCAGCCGCGCGATGAACCTCTCGGTAAATCCCAACAATCG 935  
Db 2319 CCGGAGCCCAAGAAACAAGCAGCCGCGCGATGAACCTCTCGGTAAATCCCAACAATCG 2378  
QY 936 GCTATCGGTGGCAATCTCACAGCGGTATGACCATGGCGGAAACATGCTGACAGATC 995  
Db 2379 GCTATCGGTGGCAATCTCACAGCGGTATGACCATGGCGGAAACATGCTGACAGATC 2438  
QY 996 CAGGAAAGCCAGTTCGATTCCTCAATCCGATTTGTGAGCAGCAATAGCGTGAATCAACACAGC 1055  
Db 2439 CAGGAAAGCCAGTTCGATTCCTCAATCCGATTTGTGAGCAGCAATAGCGTGAATCAACACAGC 2498  
QY 1056 AACTCCAAACACTACGCGCGGAGGTGGTGACACCAACGTCGACGACCACTGGATTA 1115  
Db 2499 AACTCCAAACACTACGCGCGGAGGTGGTGACACCAACGTCGACGACCACTGGATTA 2558  
QY 1116 CAGCGTCCAGGTACTCGCAAGCTGTACCGGCGGAGCAGCAGCAAGATCGCTCGGG 1175  
Db 2559 CAGCGTCCAGGTACTCGCAAGCTGTACCGGCGGAGCAGCAGCAAGATCGCTCGGG 2618  
QY 1176 GTGCTTAATAACACATTCGATTCCTCAATCCGAGCGGAGCAGTGTGTTCCGCGGACAA 1235  
Db 2619 GTGCTTAATAACACATTCGATTCCTCAATCCGAGCGGAGCAGTGTGTTCCGCGGACAA 2678  
QY 1236 GAGATGCCACTCGCGGCCAAACCCAGGAGCTGGCGGCTAGTACCCGCGCATCGCAC 1295  
Db 2679 GAGATGCCACTCGCGGCCAAACCCAGGAGCTGGCGGCTAGTACCCGCGCATCGCAC 2738  
QY 1296 GCATTCGAAGAGGTGGAGCGACTGAAGAGATTCGGGAGAGATCGGAGAGCGAGCGCCGACAGGCC 1355

Db 2739 GCATTCGAAGAGGTGGAGCGACTGAAGAGAGAAATCGCGAGAGCGACGCGCCGACAGGCC 2798  
QY 1356 GAGATGAAGAGGAGAGAGGTGGCGCTGATGAACACGAGATCCGCGCAATCCAACTCGGAG 1415  
Db 2799 GAGATGAAGAGGAGAGAGGTGGCGCTGATGAACACGAGATCCGCGCAATCCAACTCGGAG 2858  
QY 1416 AGCGCGCAATGATACGCGCAATATCAGAGCAGCTCGAATTTGTCGCGCTGCTCGATGGC 1475  
Db 2859 AGCGCGCAATGATACGCGCAATATCAGAGCAGCTCGAATTTGTCGCGCTGCTCGATGGC 2918  
QY 1476 CAGGCGCTGATGACCATCAGATCAGTGTGCGTGCAGAGCTCCCATTTAGCCGCAAG 1535  
Db 2919 CAGGCGCTGATGACCATCAGATCAGTGTGCGTGCAGAGCTCCCATTTAGCCGCAAG 2978  
QY 1536 GAGTCAATTCGAAGAGATCGATGTCTTTCGGTGCAGCAAGACATGCTCATCGTG 1595  
Db 2979 GAGTCAATTCGAAGAGATCGATGTCTTTCGGTGCAGCAAGACATGCTCATCGTG 3038  
QY 1596 CAGGAGCGCGCAGCAGAGTGCACCTCACCAGTTCTCGAGAACCAAGTTTCGCTTC 1655  
Db 3039 CAGGAGCGCGCAGCAGAGTGCACCTCACCAGTTCTCGAGAACCAAGTTTCGCTTC 3098  
QY 1656 GACTAGCGCTTCAACGACACGTCGACATGCTGATGATCAATACACAGCCAGCC 1714  
Db 3099 GACTAGCGCTTCAACGACACGTCGACATGCTGATGATCAATGAGTATCCAAATCC 3157

RESULT 5  
AC017670 35708 bp DNA linear HTG 10-DEC-1999  
LOCUS Drosophila melanogaster, \*\*\* SEQUENCING IN PROGRESS \*\*\*.  
DEFINITION AC017670  
AC017670.1 GI:6554327  
VERSION HTG; HTGS, PHASE2.  
KEYWORDS Drosophila melanogaster (fruit fly)  
SOURCE Drosophila melanogaster  
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 35708)  
AUTHORS Adams, M. and Venter, J.C.  
TITLE Direct Submission  
JOURNAL Submitted (10-DEC-1999) Celera Genomics, 45 West Gude Drive,  
Rockville, MD, USA  
COMMENT This sequence was identified as CDM:10211672 by the submitter.  
For more information on this record e-mail to fly@celera.com.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

FEATURES  
source  
1..35708  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"

ORIGIN  
Query Match 50.2%; Score 1547.8; DB 14; Length 35708;  
Best Local Similarity 99.6%; Pred. No. 2.5e-286;  
Matches 1552; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 156 AGCCAGCGGCTGCGGTTTTTCTTTTATCGTTGCCAGCGCAACGAAACGACA 215  
Db 2212 AGCCAGCGGCTGCGGTTTTTCTTTTATCGTTGCCAGCGCAACGAAACGACA 2271  
QY 216 GTTGCAATTTGATTCAGCACAAACACATCTAACCGCGACCCGCGAGAGACACAC 275  
Db 2272 GTTGCAATTTGATTTAGCACAAACACATCTAACCGCGACCCGCGAGCAGCACAC 2331  
QY 276 ACACACACTGGGACACTCGAAAAAACAACAGACGCTGTCGGGAGCTCGACAGAG 335  
Db 2332 ACACACACTGGGACACTCGAAAAAACAACAGACGCTGTCGGGAGCTCGACAGAG 2391





Db 69229 TACAGGTGAGTCAAAAGAGAAAGGATCATGGACATGATTACGGTGGGCGAGAGCGTCAAG 69170  
QY 696 ATCAAGCGGACGAGTGGCGCGCTCCACATGGCGGCTGGTGGCGGTGATCAACCGATCGGGC 755  
Db 69169 ATCAAGCGGACGAGTGGCGCGCTCCACATGGCGGCTGGTGGCGGTGATCAACCGATCGGGC 69110  
QY 756 AAGTGCATCAAGTCAAGTGTACGAGCGCGCGGAAACGAAGCGGCAAGAGGAGGTAGAACTG 815  
Db 69109 AAGTGCATCAAGTCAAGTGTACGAGCGCGGAAACGAAGCGGCAAGAGGAGGTAGAACTG 69050  
QY 816 GACGCCATCACTACGCTCAATCCGAGTCAATGCAAGTCAATGCTCGAACAGCACCGCCGCC 875  
Db 69049 GACGCCATCACTACGCTCAATCCGAGTCAATGCAAGTCAATGCTCGAACAGCACCGCCGCC 68990  
QY 876 CCGGAGCCCAAGAACAGCACCGCGCGGATGACCTCTCGCTAAATCCACACAACTCG 935  
Db 68989 CCGGAGCCCAAGAACAGCACCGCGCGGATGACCTCTCGCTAAATCCACACAACTCG 68930  
QY 936 GCTATCGGTGGCAATCTCACACAGCGGTATGACCATGGCGGAAACATGCTGAACAAGATC 995  
Db 68929 GCTATCGGTGGCAATCTCACACAGCGGTATGACCATGGCGGAAACATGCTGAACAAGATC 68870  
QY 996 CAGGAAAGCGATCGATTCCCAATCCGATTTGTCAGCAGCAATAGCGTGAATACAAACAGC 1055  
Db 68869 CAGGAAAGCGATCGATTCCCAATCCGATTTGTCAGCAGCAATAGCGTGAATACAAACAGC 68810  
QY 1056 AACTCCACACTACCGCGCGGAGGTGGTGCACACAACTGACACCACTGCGATTA 1115  
Db 68809 AACTCCACACTACCGCGCGGAGGTGGTGCACACAACTGACACCACTGCGATTA 68750  
QY 1116 CAGCGTCCACCGTACTCGCAAGCTCTACCGGCCAGCAGCAGCAAGGATCGCTCGCGC 1175  
Db 68749 CAGCGTCCACCGTACTCGCAAGCTCTACCGGCCAGCAGCAGCAAGGATCGCTCGCGC 68690  
QY 1176 GTGCTTAATACATTCGCAATCCCAAGCGCGGAGCGTGTGCTCGCGCGCACAA 1235  
Db 68689 GTGCTTAATACATTCGCAATCCCAAGCGCGGAGCGTGTGCTCGCGCGCACAA 68630  
QY 1236 GAGTTCGCACTCGCGCCACACACCGAGGAGTGGCGGCTGTACCGCGGATCGCAC 1295  
Db 68629 GAGTTCGCACTCGCGCCACACACCGAGGAGTGGCGGCTGTACCGCGGATCGCAC 68570  
QY 1296 GCATTGAAGAGGTGGAGCGACTGAAGGAGATCCGAGAGCGACGCGCGCACAGGCC 1355  
Db 68569 GCATTGAAGAGGTGGAGCGACTGAAGGAGATCCGAGAGCGACGCGCGCACAGGCC 68510  
QY 1356 GAGTGAAGAGGAGGAGGTCGCTGATGAACCGAGATCCGCGGCAATCCAAATCGGAG 1415  
Db 68509 GAGTGAAGAGGAGGAGGTCGCTGATGAACCGAGATCCGCGGCAATCCAAATCGGAG 68450  
QY 1416 ACGGCGCAATGATACGCGAATATCAGAGCAGCTGGATTTGTGCGGCTCTCGATGGC 1475  
Db 68449 ACGGCGCAATGATACGCGAATATCAGAGCAGCTGGATTTGTGCGGCTCTCGATGGC 68390  
QY 1476 CAGGCGGTCGATGACCATCAGATCAGTGTGCGTGGCGCAAGCGTCCCATAGCGCGAAG 1535  
Db 68389 CAGGCGGTCGATGACCATCAGATCAGTGTGCGTGGCGCAAGCGTCCCATAGCGCGAAG 68330  
QY 1536 GAGTCAATTCGAGGAGATCGATGCTATTTGCGTGGCGCGCAAGGACATGCTCATCGTG 1595  
Db 68329 GAGTCAATTCGAGGAGATCGATGCTATTTGCGTGGCGCGCAAGGACATGCTCATCGTG 68270  
QY 1596 CACGAGCGCGCAGGAGGTGACCTACCAAGTTCCTGGAGAACCAAGATTTGCGTTTC 1655  
Db 68269 CACGAGCGCGCAGGAGGTGACCTACCAAGTTCCTGGAGAACCAAGATTTGCGTTTC 68210  
QY 1656 GACTACGCGCTTCAACGACACGTCGACAAATGCGATGCTATCAATACAAATACAGAGCAAGCC 1714  
Db 68209 GACTACGCGCTTCAACGACACGTCGACAAATGCGATGCTATCAATACAAATACAGAGCAAGCC 68151

RESULT

AE003485/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AE003485 283075 bp DNA linear INV 10-AUG-2005  
Drosophila melanogaster chromosome X, section 37 of the  
complete sequence.  
AE003485 AE002593 AE014298  
AE003485.3 GI:22832075  
Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pelecygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE  
AUTHORS

1 (bases 1 to 283075)  
Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
Brandon R.C., Rogers Y.H., Blaise J., Champs M., Pfeiffer B.D.,  
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
Abril J.F., Agbayani A., An H.J., Andrews Pannkock C., Baldwin D.,  
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattner P.,  
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A.,  
Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B.,  
Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I.,  
Dietz S.M., Dodson K., Doup L.E., Downes M., Dugan-Rocha S.,  
Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C.,  
Ferrier S., Fleischmann W., Fowler C., Gabriellian A.E., Garg N.S.,  
Gelbart W.M., Glasser K., Glodek A., Gong P., Gorrell J.H., Gu Z.,  
Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J.,  
Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J.,  
Wei M.H., Ibegwam C., Jalali M., Kalush F., Karpen G.H., Ke Z.,  
Kernison J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C.,  
Kravitz S., Kulp D., Lai Z., Laok B., Lei Y., Levitsky A.A.,  
Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C.,  
McLeod M.P., McPherson D., Merkulov G., Mitsuhashi N.V., Mobarry C.,  
Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L.,  
Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K.,  
Nusker D.R., Paclet J.M., Palazzolo M., Pittman G.S., Pan S.,  
Pollard J., Puri V., Reese M.G., Reinert K., Remington K.,  
Saunders R.D., Scheeler F., Shen H., Shue B.C., Siden-Kiamos I.,  
Simpson M., Skupski M.P., Smith T., Spier E., Spradling A.C.,  
Stapleton M., Strong R., Sun E., Svirskas R., Tector C., Turner R.,  
Venter E., Wang A.H., Wang X., Wang Z.Y., Wasearman D.A.,  
Weinstock G.M., Weissbach J., Williams S.M., Woodage J.,  
Worley K.C., Wu D., Yang S., Yao Q.A., Ye J., Ye R.F.,  
Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Zheng X.H.,  
Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
Gibbs R.A., Myers E.W., Rubin G.M. and Venter J.C.  
The genome sequence of Drosophila melanogaster  
Science 287 (5461), 2185-2195 (2000)

TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS

2 (bases 1 to 283075)  
Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W.,  
Halpern A., Patel S., Adams M., Champe M., Dugan S.P., Friese E.,  
Hodgson A., George R.A., Hoskins R.A., Lavery T., Muzny D.M.,  
Nelson C.R., Paclet J.M., Park S., Pfeiffer B.D., Richards S.,  
Sodergren E.J., Svirskas R., Tabor P.E., Wan K., Stapleton M.,  
Sutton G.G., Venter C., Weinstock G., Scherer S.E., Myers E.W.,  
Gibbs R.A. and Rubin G.M.  
Finishing a whole-genome shotgun: release 3 of the Drosophila  
melanogaster euchromatic genome sequence  
Genome Biol. 3 (12), RESEARCH0079 (2002)

TITLE  
JOURNAL  
PUBMED  
REFERENCE  
AUTHORS

3 (bases 1 to 283075)  
Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L.,  
Berman B.P., Bettencourt B.R., Celniker S.E., de Grey A.D.,  
Drysdale R.A., Harris N.L., Richter J., Russo S., Schroeder A.J.,  
Shu S.Q., Stapleton M., Yamada C., Ashburner M., Gelbart W.M.,  
Rubin G.M. and Lewis S.E.

**TITLE** Annotation of the Drosophila melanogaster euchromatic genome: a systematic review

**JOURNAL** Genome Biol. 3 (12), RESEARCH0083 (2002)

**PUBMED** 12537572

**REFERENCE** 4 (bases 1 to 283075)

**AUTHORS** Kaminker, J.S., Bergman, C.M., Kronmiller, B., Carlson, J., Svirekas, R., Patel, S., Frise, E., Wheeler, D.A., Lewis, S.E., Rubin, G.M., Ashburner, M. and Celnikier, S.E.

**TITLE** The transposable elements of the Drosophila melanogaster euchromatin: a genomics perspective

**JOURNAL** Genome Biol. 3 (12), RESEARCH0084 (2002)

**PUBMED** 12537573

**REFERENCE** 5 (bases 1 to 283075)

**AUTHORS** Celnikier, S., Carlson, J., Wan, K., Pfeiffer, B., Frise, E., George, R., Hoskins, R., Stapleton, M., Pacleb, J., Park, S., Svirekas, R., Smith, E., Yu, C. and Rubin, G.

**CONSRM** Berkeley Drosophila Genome Project

**TITLE** Drosophila melanogaster release 4 sequence

**JOURNAL** Unpublished

**REFERENCE** 6 (bases 1 to 283075)

**AUTHORS** Adams, M.D., Celnikier, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.

**TITLE** Direct Submission

**JOURNAL** Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

**REFERENCE** 7 (bases 1 to 283075)

**AUTHORS** FlyBase

**CONSRM** Direct Submission

**TITLE** Submitted (22-JUL-2005) FlyBase, Harvard University, Biological Laboratories, 16 Divinity Ave, Cambridge, MA 02138, USA

**JOURNAL** On Sep 13, 2002 this sequence version replaced gi:10728182.

**COMMENT** Location/Qualifiers

**FEATURES**

- 1. 283075
  - /organism="Drosophila melanogaster"
  - /mol\_type="genomic DNA"
  - /db\_xref="taxon:7227"
  - /chromosome="X"
- 2769..3589
  - /gene="CG2081"
  - /locus\_tag="Dmel\_CG2081"
  - /old\_locus\_tag="CG2081"
  - /map="10A4-10A4"
  - /db\_xref="FLYBASE:FBgn0030262"
- Join(2769..2931,3000..3138,3205..3589)
  - /gene="CG2081"
  - /locus\_tag="Dmel\_CG2081"
  - /old\_locus\_tag="CG2081"
  - /product="CG2081-RE"
  - /db\_xref="FLYBASE:FBgn0030262"
- Join(2850..2931,3000..3138,3205..3466)
  - /gene="CG2081"
  - /locus\_tag="Dmel\_CG2081"
  - /old\_locus\_tag="CG2081"
  - /note="CG2081 gene product from transcript CG2081-RE"
  - /codon\_start=1
  - /product="CG2081-PB"
  - /protein\_id="AAF47993.1"
  - /db\_xref="GI:7292594"
  - /db\_xref="FLYBASE:FBgn0030262"
  - /translation="MESISSMILVAMSLIIGSSQATPYRPSAVLYNQVCMDFLTGRLYIGEVTRDQCVRIOCLTQLWEDSCQVPLTQNCPTPVESTNPHAYPRCCPLYEKSYSNSGGTLEQTNIYDHYGLRSLSHLTEMIVDGRTPPRGEIHTSARKYQV"
- 3860..5573
  - /gene="CG2076"
  - /locus\_tag="Dmel\_CG2076"
  - /old\_locus\_tag="CG2076"
  - /map="10A4-10A4"
  - /db\_xref="FLYBASE:FBgn0030263"
  - Join(3860..4277,4345..4456,4530..4671,4770..5106,5204..5336,5400..5573)
    - /gene="CG2076"
    - /locus\_tag="Dmel\_CG2076"

**CDS**

- /old\_locus\_tag="CG2076"
- /product="CG2076-RA"
- /db\_xref="FLYBASE:FBgn0030263"
- Join(4058..4277,4345..4456,4530..4671,4770..5106,5204..5336,5400..5481)
  - /gene="CG2076"
  - /locus\_tag="Dmel\_CG2076"
  - /old\_locus\_tag="CG2076"
  - /note="CG2076 gene product from transcript CG2076-RA"
  - /codon\_start=1
  - /product="CG2076-PA"
  - /protein\_id="AAF47994.1"
  - /db\_xref="GI:7292595"
  - /translation="MLRLALSTAPAGAKILAQSPQLFRANPNLNLSTVRTYA RRVTRTPPEEMRAPSLKEMLGPPSNAYSMGKAAAGAAVGLGALCYGVGLGK QFSIADNAIMWPFVRDRITQSTVAFPGSCVLTAATAATFESHRLLELASGGILAT IASLALVIGSGAVARSIEYQPLGAKHLAWAHCALGNVIAPICFPMGSPILTRALY TCGIVGGLSTIAACAPSDKFLYMGGLAIGLVVFASSLASMWLPTTALGAGLASMS LYGGJLVLFSGFLLYDTRMVRRAEVPYQSYTPYPFINASMSIYMDVLNIFIRIVTIL SGGQRRK"
  - 5972..8143
    - /gene="CG2061"
    - /locus\_tag="Dmel\_CG2061"
    - /old\_locus\_tag="CG2061"
    - /map="10A4-10A4"
    - /db\_xref="FLYBASE:FBgn0027498"
    - Join(5972..6197,6260..8143)
      - /gene="CG2061"
      - /locus\_tag="Dmel\_CG2061"
      - /old\_locus\_tag="CG2061"
      - /product="CG2061-RC, transcript variant C"
      - /db\_xref="FLYBASE:FBgn0027498"
    - Join(5972..6197,6260..6320,6438..8143)
      - /gene="CG2061"
      - /locus\_tag="Dmel\_CG2061"
      - /old\_locus\_tag="CG2061"
      - /product="CG2061-RA, transcript variant A"
      - /db\_xref="FLYBASE:FBgn0027498"
    - Join(6208..6320,6438..8143)
      - /gene="CG2061"
      - /locus\_tag="Dmel\_CG2061"
      - /old\_locus\_tag="CG2061"
      - /product="CG2061-RB, transcript variant B"
      - /db\_xref="FLYBASE:FBgn0027498"
    - 6524..7783
      - /gene="CG2061"
      - /locus\_tag="Dmel\_CG2061"
      - /old\_locus\_tag="CG2061"
      - /product="CG2061 gene product from transcript CG2061-RA"
      - /codon\_start=1
      - /product="CG2061-PA, isoform A"
      - /protein\_id="AAF47995.1"
      - /db\_xref="GI:7292596"
      - /db\_xref="FLYBASE:FBgn0027498"

**CDS**

- Query Match 50.2%; Score 1547.8; DB 2; Length 283075;
- Best Local Similarity 99.6%; Pred. No. 2.8e-286;
- Matches 1552; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

- QY 156 AGCCAGCGGCGTCGGTTCCTTTTCTTTTATCCGTGGCAGACGAAACGAAACGACA 215
- DB 49902 AGCCAGCGGCGTCGGTTCCTTTTCTTTTATCCGTGGCAGACGAAACGACA 49843
- QY 216 GTTGGCATTTGAATTCAGCAACACACATACATACCGCCGACCCGAGCAGCACAC 275
- DB 49842 GTTGGCATTTGAATTCAGCAACACACATACATACCGCCGACCCGAGCAGCACAC 49783
- QY 276 ACACACACTGGGACACTCGAAAAAAGAAAAACAGACGCTGTCCGGCAGCTCCACAAGCAG 335
- DB 49782 ACACACACTGGGACACTCGAAAAAAGAAAAACAGACGCTGTCCGGCAGCTCCACAAGCAG 49723
- QY 336 TTGGGTTTCGATTTAGTTGTCATTCGCTTGAATTCGGTTCGGGGGCTTAGTTTCCACAAGTT 395



49722	TTGGGTTTCGATTTAGTTGTCAATGCTTGAATTCGGTTCGGGCTTAGTTTCCACAAGTT	49663	Db
396	TATCGCTCGTCAAGAAACAACAAATATAAATTTATTTTCGACCTAAAAATCTGACTAAAT	455	Qy
49662	TATCGCTCGTCAAGAAACAACAAATATAAATTTATTTTCGACCTAAAAATCTGACTAAAT	49603	Db
456	TGTGTTTTTTTGTATTATTTATTTAGGCACATTTTGACACCAACAACGTAGTTACTAC	515	Qy
49602	TGTGTTTTTTTGTATTATTTATTTAGGCACATTTTGACACCAACAACGTAGTTACTAC	49543	Db
516	ATCTAGCACTAACGGAATCTCTCTGCAAGCAGTGGAAAGTTGCTGTCCATCAAGCAGTAC	575	Qy
49542	ATCTAGCACTAACGGAATCTCTCTGCAAGCAGTGGAAAGTTGCTGTCCATCAAGCAGTAC	49483	Db
576	TCGGAGTTTAAACCGAGTAAGCCGGAGAAAGAGAGATCGGTGGAGATAGAGATA	635	Qy
49482	TCGGAGTTTAAACCGAGTAAGCCGGAGAAAGAGAGATCGGTGGAGATAGAGATA	49423	Db
636	TACAGTGGAGTCAAGAGAAAGGATCATGGACATGATTACGGTGGGGCAGAGCGTCAAG	695	Qy
49422	TACAGTGGAGTCAAGAGAAAGGATCATGGACATGATTACGGTGGGGCAGAGCGTCAAG	49363	Db
696	ATCAAGCGGAACGGATGGCGCGTCCACATGGCCGTGGTGGCGGTGATCAACCCAGTCGGGC	755	Qy
49362	ATCAAGCGGAACGGATGGCGCGTCCACATGGCCGTGGTGGCGGTGATCAACCCAGTCGGGC	49303	Db
756	AGTGCATCACAGTCCGATGGTACAGCGCGCGGAAACGAAGGGCAAGAGGTAGAACTG	815	Qy
49302	AGTGCATCACAGTCCGATGGTACAGCGCGCGGAAACGAAGGGCAAGAGGTAGAACTG	49243	Db
816	GACGCCATACCTCACGCTCAATCCGAGCTAATGCAAGATACTGTGCAACAGCAGCCGCC	875	Qy
49242	GACGCCATACCTCACGCTCAATCCGAGCTAATGCAAGATACTGTGCAACAGCAGCCGCC	49183	Db
876	CCGGAGCCCAAGAAACAAGCCACCGCGCGATGAACCTCTCGGTAATCCCAACAATCG	935	Qy
49182	CCGGAGCCCAAGAAACAAGCCACCGCGCGATGAACCTCTCGGTAATCCCAACAATCG	49123	Db
936	GCTATCGGTGCAATCTCACCAGCGGTATGACCATGGCGGGAACATCTGAACAGATC	995	Qy
49122	GCTATCGGTGCAATCTCACCAGCGGTATGACCATGGCGGGAACATCTGAACAGATC	49063	Db
996	CAGGAAAGCCAGTCCGATTTCCCAATCCGATTTGTGACAGCAATAGGTTGAATACAAACAGC	1055	Qy
49062	CAGGAAAGCCAGTCCGATTTCCCAATCCGATTTGTGACAGCAATAGGTTGAATACAAACAGC	49003	Db
1056	AACTCCAACTACCGCGCGGAGGTGGTGGCAACCAACGTGCAACCACTGGATTA	1115	Qy
49002	AACTCCAACTACCGCGCGGAGGTGGTGGCAACCAACGTGCAACCACTGGATTA	48943	Db
1116	CAGGTCCACCGTACTCCGACAGTCTACCGCGCAGCAGCAGAGATCGCTCGCG	1175	Qy
48942	CAGGTCCACCGTACTCCGACAGTCTACCGCGCAGCAGCAGAGATCGCTCGCG	48883	Db
1176	GTGCTTAATAACACATTCGCCAATCCACGCGCGCAGCAGTGTGTCGGCGGCACAA	1235	Qy
48882	GTGCTTAATAACACATTCGCCAATCCACGCGCGCAGCAGTGTGTCGGCGGCACAA	48823	Db
1236	GGAGTCCGCACTGCGGCGCAACCAACGAGGAGTGGCGCGGTAGTACCCCGCGATCGCAC	1295	Qy
48822	GGAGTCCGCACTGCGGCGCAACCAACGAGGAGTGGCGCGGTAGTACCCCGCGATCGCAC	48763	Db
1296	GCATTGAAGAGGTGGAGCGACTGAGGAGATCGCGAGAGCGACGCGCGCAGCGCC	1355	Qy
48762	GCATTGAAGAGGTGGAGCGACTGAGGAGATCGCGAGAGCGACGCGCGCAGCGCC	48703	Db
1356	GAGATGAAGAGGAGAGGTGGCGTGTATGAACAGAGTCCGGGCAATCCAAACTGGGAG	1415	Qy
48702	GAGATGAAGAGGAGAGGTGGCGTGTATGAACAGAGTCCGGGCAATCCAAACTGGGAG	48643	Db
1416	ACGGCGCAATGATACCGCAATATACAGACACGCTGGAAATTTGTGCGCGCTCTCGATGGC	1475	Qy
48642	ACGGCGCAATGATACCGCAATATACAGACACGCTGGAAATTTGTGCGCGCTCTCGATGGC	48583	Db
1476	CAGCGCGTCCGATGACCATCAGATCACAGTGTGCGTGGCGCAAGCGTCCCATTAGCGCAAG	1535	Qy
48582	CAGCGCGTCCGATGACCATCAGATCACAGTGTGCGTGGCGCAAGCGTCCCATTAGCGCAAG	48523	Db
1536	GAGGTCAATCCAAAGGAGATCGATGTCTATTTCGGTGGCGGCAAGGACATGCTCATCGTG	1595	Qy
48522	GAGGTCAATCCAAAGGAGATCGATGTCTATTTCGGTGGCGGCAAGGACATGCTCATCGTG	48463	Db
1596	CACGAGCGCGCACCAAGGTGACCTCACCAAGTTCCTGGAGAACCAAGTTCGCTTC	1655	Qy
48462	CACGAGCGCGCACCAAGGTGACCTCACCAAGTTCCTGGAGAACCAAGTTCGCTTC	48403	Db
1656	GACTACGCTTCACGACACGTCGACATGCCATGCTGTATCAATATACAGCCCAAGCC	1714	Qy
48402	GACTACGCTTCACGACACGTCGACATGCCATGCTGTATCAATATACAGCCCAAGCC	48344	Db
RESULT 8			
AC116537			
LOCUS			
DEFINITION	Drosophila melanogaster chromosome X clone RP98-9f14, WORKING DRAFT		
SEQUENCE	14 unordered pieces.		
ACCESSION	AC116537		
VERSION	AC116537.2 GI:20898595		
KEYWORDS	HTG; HTGS PHASE1; HTGS DRAFT; HTGS ACTIVEPIN.		
SOURCE	Drosophila melanogaster (fruit fly)		
ORGANISM	Drosophila melanogaster		
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
AUTHORS	1 (bases 1 to 334028)		
	Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alsbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T., Barbacia, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, B., Homsi, P., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseg, H., Lozado, R.X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Maxiney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkenwo, S., Oguh, M., Okwionu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pichens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonaika, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Umari, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G. and Gibbs, R.		

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Direct Submission  
Unpublished  
2 (bases 1 to 334028)  
Worley, K.C.  
Direct Submission  
Submitted (29-MAR-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 334028)  
Worley, K.C.  
Direct Submission  
Submitted (18-MAY-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On May 17, 2002 this sequence version replaced gi:19807669.

COMMENT

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: DRIT  
Center clone name: RP98-9F14  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 327497 bases at least Q40  
Consensus quality: 328682 bases at least Q30  
Consensus quality: 329600 bases at least Q20  
Estimated insert size: 175224; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/Genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 14 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 2752: contig of 2752 bp in length  
\* 2753 2852: gap of unknown length  
\* 2853 9484: contig of 6632 bp in length  
\* 9484 9584: gap of unknown length  
\* 9585 13752: contig of 4168 bp in length  
\* 13753 13852: gap of unknown length  
\* 13853 30299: contig of 16447 bp in length  
\* 30300 30399: gap of unknown length  
\* 30400 45744: contig of 15345 bp in length  
\* 45745 45844: gap of unknown length  
\* 45845 59632: contig of 13788 bp in length  
\* 59633 59732: gap of unknown length  
\* 59733 80931: contig of 21199 bp in length  
\* 80932 81031: gap of unknown length  
\* 81032 104548: contig of 23517 bp in length  
\* 104549 104648: gap of unknown length  
\* 104649 123841: contig of 19193 bp in length  
\* 123842 123941: gap of unknown length  
\* 123942 145561: contig of 21620 bp in length  
\* 145562 145661: gap of unknown length  
\* 145662 183677: contig of 38016 bp in length  
\* 183678 183777: gap of unknown length  
\* 183778 217275: contig of 33498 bp in length  
\* 217276 217375: gap of unknown length  
\* 217376 268427: contig of 51052 bp in length  
\* 268428 268527: gap of unknown length  
\* 268528 334028: contig of 65501 bp in length.

FEATURES  
source  
1. .334028

/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/chromosome="X"  
/clone="RP98-9F14"

2753..2852  
/estimated\_length=unknown  
9485..9584  
/estimated\_length=unknown  
13753..13852  
/estimated\_length=unknown  
30300..30399  
/estimated\_length=unknown  
45745..45844  
/estimated\_length=unknown  
59633..59732  
/estimated\_length=unknown  
80932..81031  
/estimated\_length=unknown  
104549..104648  
/estimated\_length=unknown  
123842..123941  
/estimated\_length=unknown  
145562..145661  
/estimated\_length=unknown  
183678..183777  
/estimated\_length=unknown  
217276..217375  
/estimated\_length=unknown  
268428..268527  
/estimated\_length=unknown

ORIGIN

Query Match 50.2%; Score 1547.8; DB 14; Length 334028;  
Best Local Similarity 99.6%; Pred. No. 2.8e-286; Indels 0; Gaps 0;  
Matches 1552; Conservative 0; Mismatches 7;  
QY 156 AGCCAGCGGCTGCGGTTTTCCTTTTGTATCGGTTGCCAGAGCGCAACGAAACGACA 215  
Db 217943 AGCCAGCGGCTGCGGTTTTCCTTTTGTATCGGTTGCCAGAGCGCAACGAAACGACA 218002  
QY 216 GTTGGCATTTGAATTCAGCAAAACACACATACATACGCGGACCGCGAGCAGCACAC 275  
Db 218003 GTTGGCATTTGAATTCAGCAAAACACACATACATACGCGGACCGCGAGCAGCACAC 218062  
QY 276 ACACACACTGGGACACTCGAAAAAAGAAAAACAGACGCTGTCGGGACCTCCGACAGCAG 335  
Db 218063 ACACACACTGGGACACTCGAAAAAAGAAAAACAGACGCTGTCGGGACCTCCGACAGCAG 218122  
QY 336 TTGGGTTTCGATTTAGTTTGTCAATGCTTGAATTCGGTTTCGGGGCTTAGTTTCCACAAGTT 395  
Db 218123 TTGGGTTTCGATTTAGTTTGTCAATGCTTGAATTCGGTTTCGGGGCTTAGTTTCCACAAGTT 218182  
QY 396 TATCGCTCGTCAAGAAACAAACGAAATAAATTTATTTTCGACCTAAATAAATCTGACTAAAT 455  
Db 218183 TATCGCTCGTCAAGAAACAAACGAAATAAATTTATTTTCGACCTAAATAAATCTGACTAAAT 218242  
QY 456 TGTGTTTTTTTGTATGTTATTTATTTAGGCACATTTTCGACACCAACGAGTGTACTACTAC 515  
Db 218243 TGTGTTTTTTTGTATGTTATTTATTTAGGCACATTTTCGACACCAACGAGTGTACTACTAC 218302  
QY 516 ATCTACGACTAACCGAACTCCTCTCGAAGCAGTGGAAAGTTGCTGTCCATCAAGCAGTAC 575  
Db 218303 ATCTACGACTAACCGAACTCCTCTCGAAGCAGTGGAAAGTTGCTGTCCATCAAGCAGTAC 218362  
QY 576 TCGGAGTTAAACGACGATAAGCCGGGAGAAAGAGAAAGAGATCGGTGGAGAAATAGAGATA 635  
Db 218363 TCGGAGTTAAACGACGATAAGCCGGGAGAAAGAGAAAGAGATCGGTGGAGAAATAGAGATA 218422  
QY 636 TACAGTGGAGTCAAGAGGAGGAGATCATGACATGTTTACGTTGGGCGAGCGCTCAAG 695  
Db 218423 TACAGTGGAGTCAAGAGGAGGAGATCATGACATGTTTACGTTGGGCGAGCGCTCAAG 218482

```
QY 696 ATCAAGCGGACGAGTGGCGCGCTCCACATGGCGGTGGTGGCGGTGATCAACCAAGTCGGGC 755
Db 218483 ATCAAGCGGACGAGTGGCGCGCTCCACATGGCGGTGGTGGCGGTGATCAACCAAGTCGGGC 218542
QY 756 AAGTGCATCAGTGCAGTGTACGAGCGCGCGGAAACGAAGGGCAAGAGGTAGAACTG 815
Db 218543 AAGTGCATCAGTGCAGTGTACGAGCGCGCGGAAACGAAGGGCAAGAGGTAGAACTG 218602
QY 816 GAGCGCATCTACGCTCAATCCGAGCTAATGCAAGTACTGTCGACAGCAGCCCGCC 875
Db 218603 GAGCGCATCTACGCTCAATCCGAGCTAATGCAAGTACTGTCGACAGCAGCCCGCC 218662
QY 876 CGGAGCGCCCAAGAAACAGCCACCGCGCGATGAACCTCTCCGCTAAATCCCAACAACTCG 935
Db 218663 CGGAGCGCCCAAGAAACAGCCACCGCGCGATGAACCTCTCCGCTAAATCCCAACAACTCG 218722
QY 936 GTATCGGTGGCAATCTACAGCCGCTATGACCATGGCGGAAACATGCTGAACAAATC 995
Db 218723 GTATCGGTGGCAATCTACAGCCGCTATGACCATGGCGGAAACATGCTGAACAAATC 218782
QY 996 CAGGAAAGCCAGTGCATTCCTCAATCCGATTTGTCAGCAGCAATAGCTGAATACAAACAGC 1055
Db 218783 CAGGAAAGCCAGTGCATTCCTCAATCCGATTTGTCAGCAGCAATAGCTGAATACAAACAGC 218842
QY 1056 AACTCCAACTACGCGCGCGGAGGTGGTGACCAACAACTCGACCACTCGATTA 1115
Db 218843 AACTCCAACTACGCGCGCGGAGGTGGTGACCAACAACTCGACCACTCGATTA 218902
QY 1116 CAGCTTCACGCTACTCGCAAGCTCTACCGGCGAGCAGCAGCAAGGATCGCTCGCGC 1175
Db 218903 CAGCTTCACGCTACTCGCAAGCTCTACCGGCGAGCAGCAGCAAGGATCGCTCGCGC 218962
QY 1176 GTGCTTAATAACACTTCCCAATCCAGCGCGGAGCGAGTGTGCTGGCGCGCACAA 1235
Db 218963 GTGCTTAATAACACTTCCCAATCCAGCGCGGAGCGAGTGTGCTGGCGCGCACAA 219022
QY 1236 GGAGTCGCCACTCGCGCCCAACACCGAGGAGTGGCGCGGTAGTACCCGCGCATCGAC 1295
Db 219023 GGAGTCGCCACTCGCGCCCAACACCGAGGAGTGGCGCGGTAGTACCCGCGCATCGAC 219082
QY 1296 GCATTGAAAGAGTGGCGGACTGAAGGAGAAATCGCGAAGCGCAGCGCCGCGCACAGGCC 1355
Db 219083 GCATTGAAAGAGTGGCGGACTGAAGGAGAAATCGCGAAGCGCAGCGCCGCGCACAGGCC 219142
QY 1356 GAGATGAAGGAGGAGAGTGGCTGATGAACAGGATCCGGCAATCCAAATCGGAG 1415
Db 219143 GAGATGAAGGAGGAGAGTGGCTGATGAACAGGATCCGGCAATCCAAATCGGAG 219202
QY 1416 ACGCGGCAATGATACGCGAAATATCAGAGCAGCTGGAAATTTGTGCCGCTGCTCGATGGC 1475
Db 219203 ACGCGGCAATGATACGCGAAATATCAGAGCAGCTGGAAATTTGTGCCGCTGCTCGATGGC 219262
QY 1476 CAGGCGCTGATGACCATCAGATCAGTGTGGTGGCGCAAGCGTCCCATTTAGCGCAAG 1535
Db 219263 CAGGCGCTGATGACCATCAGATCAGTGTGGTGGCGCAAGCGTCCCATTTAGCGCAAG 219322
QY 1536 GAGGTCAATCGCAGGAGATCGATGCTATTTCGGTGGCGCGCAGGACATGCTCATCGTG 1595
Db 219323 GAGGTCAATCGCAGGAGATCGATGCTATTTCGGTGGCGCGCAGGACATGCTCATCGTG 219382
QY 1596 CACGAGCGCGCAGCAGGTGCAGCTCACCAAGTTCTGGAGAACCAAGTTTCGTTTC 1655
Db 219383 CACGAGCGCGCAGCAGGTGCAGCTCACCAAGTTCTGGAGAACCAAGTTTCGTTTC 219442
QY 1656 GACTACGCTTCAAGCAGCAGTGGGCAATGCCATGGTATACAAATACAGCGCAAGCC 1714
Db 219443 GACTACGCTTCAAGCAGCAGTGGGCAATGCCATGGTATACAAATACAGCGCAAGCC 219501
RESULT 8
AC007521
LOCUS 180919 bp DNA linear HTG 15-MAR-2002
DEFINITION Drosophila melanogaster strain y; cn bw sp chromosome X clone
```

```
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

TITLE
JOURNAL
REFERENCE
AUTHORS

TITLE
JOURNAL
COMMENT
```

BACR49A04 (D698) RPCI-98 49 A.4 map 10A2-10B2, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 61 unordered pieces.  
AC007521 GI:19483792  
HTG; HTGS PHASE1.  
Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
1 (bases 1 to 180919)  
Celinker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Richards,S., Sethi,H., Svirskas,R.R., Wan,K.H., Webster,D., Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.  
Sequencing of Drosophila melanogaster  
Unpublished

2 (bases 1 to 180919)  
Celinker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E., Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.  
Direct Submission  
Submitted (10-MAY-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA  
3 (bases 1 to 180919)

Celinker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Richards,S., Sethi,H., Svirskas,R.R., Wan,K.H., Webster,D., Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.  
Direct Submission  
Submitted (15-MAR-2002) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA  
On Mar 15, 2002 this sequence version replaced gi:7191020.  
For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to [bdgp@fruitfly.berkeley.edu](mailto:bdgp@fruitfly.berkeley.edu). All contigs in this submission meet the following cutoffs: length >= 200 bases.

\* NOTE: This is a 'working draft' sequence. It currently consists of 61 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 717: contig of 717 bp in length  
\* 718 797: gap of unknown length  
\* 798 1345: contig of 548 bp in length  
\* 1346 1425: gap of unknown length  
\* 1426 2041: contig of 616 bp in length  
\* 2042 2685: contig of 564 bp in length  
\* 2686 3508: contig of 743 bp in length  
\* 3509 3588: gap of unknown length  
\* 3589 4723: contig of 1135 bp in length  
\* 4724 4804: gap of unknown length  
\* 4804 5781: contig of 978 bp in length



QY	396	TATCGCTCGTCAAGAAACAAACGAAATAAAATTTATTTTCGACCTAAAAATCTGACTAAAT	455
Db	110784	TATCGCTCGTCAAGAAACAAACGAAATAAAATTTATTTTCGACCTAAAAATCTGACTAAAT	110843
QY	456	TGTGTTTTTTTGTATGTATTTATTTAGGCACATTTTGCACACACACGCTAGTTACTTAC	515
Db	110844	TGTGTTTTTTTGTATGTATTTATTTAGGCACATTTTGCACACACACGCTAGTTACTTAC	110903
QY	516	ATCTACGACTAACCGAACTCTCTCTGCAAGCAGTGGAAATGCTGCTCCATCAAGCAGTAC	575
Db	110904	ATCTACGACTAACCGAACTCTCTCTGCAAGCAGTGGAAATGCTGCTCCATCAAGCAGTAC	110963
QY	576	TCGGAGTTAAACGACGATTAAGCCGGGAGAAAGAGAAAGATTCGGTGAGAAATAGAGATA	635
Db	110964	TCGGAGTTAAACGACGATTAAGCCGGGAGAAAGAGAAAGATTCGGTGAGAAATAGAGATA	111023
QY	636	TACAGTGGAGTCAAGAGGAAGATCATGACATGATTACGTTGGGCGCAGCGTCAAG	695
Db	111024	TACAGTGGAGTCAAGAGGAAGATCATGACATGATTACGTTGGGCGCAGCGTCAAG	111083
QY	696	ATCAAGCGGACGATGCGCGCTCCACATGCGCGTGGTGGCGGTGATCAACCAAGTCGGGC	755
Db	111084	ATCAAGCGGACGATGCGCGCTCCACATGCGCGTGGTGGCGGTGATCAACCAAGTCGGGC	111143
QY	756	AAGTGCATCACAGTCGAATGGTACGAGCGCGCGGAAACGAAGGGCAAGAGGTAGAACTG	815
Db	111144	AAGTGCATCACAGTCGAATGGTACGAGCGCGCGGAAACGAAGGGCAAGAGGTAGAACTG	111203
QY	816	GACGCCATCTACGCTCAATCCGAGGTATGCAAGATATCTGTCGAACAGCAGCGCGCC	875
Db	111204	GACGCCATCTACGCTCAATCCGAGGTATGCAAGATATCTGTCGAACAGCAGCGCGCC	111263
QY	876	CGGAGGCCAAGAAACAGCCACCGCGCGATGACCTCTCGCTTAATCCACACATCG	935
Db	111264	CGGAGGCCAAGAAACAGCCACCGCGCGATGACCTCTCGCTTAATCCACACATCG	111323
QY	936	GCTATCGGTGGCAATCTCACCAGCGGTATGACCATGCGCGGAAACATGCTGAACAAGATC	995
Db	111324	GCTATCGGTGGCAATCTCACCAGCGGTATGACCATGCGCGGAAACATGCTGAACAAGATC	111383
QY	996	CAGGAAGCCAGTGTATCCCAATCCGATTTGTGACGAGCAATAGCGTGAATACAAACAGC	1055
Db	111384	CAGGAAGCCAGTGTATCCCAATCCGATTTGTGACGAGCAATAGCGTGAATACAAACAGC	111443
QY	1056	AACTCCAACTACGCGCGGAGGTGTCGACCAACAGTTCGACGACCACTGGATTA	1115
Db	111444	AACTCCAACTACGCGCGGAGGTGTCGACCAACAGTTCGACGACCACTGGATTA	111503
QY	1116	CAGCGTCCACGGTACTCGCAAGCTGCTACCGCGCCAGCAGACAAAGGATCGCTCGGCG	1175
Db	111504	CAGCGTCCACGGTACTCGCAAGCTGCTACCGCGCCAGCAGACAAAGGATCGCTCGGCG	111563
QY	1176	GTGCTTAATACATTTGCCAATCCACAGCGCGGAGCAAGT-GCTGTCGGCGGCACA	1234
Db	111564	GTGCTTAATACATTTGCCAATCCACAGCGCGGAGCAAGT-GCTGTCGGCGGCACA	111623
QY	1235	AGGAGTCCCACTCGGCGCAACCCAGGAGCTGGCGGCGCTAGTACCCGCGCATCGCA	1294
Db	111624	AGGAGTCCCACTCGGCGCAACCCAGGAGCTGGCGGCGCTAGTACCCGCGCGNNNNN	111683
QY	1295	CGCATTTGAAAGAGGTGGAGCGACTGAAGGAAATCGCGAGAAAGCGCGCCGACAGGC	1354
Db	111684	NN	111743
QY	1355	CGAGATGAGGAGAG 1370	
Db	111744	NN	111759

RESULT 10  
AC007521/c  
LOCUS 180919 bp DNA linear HTG 15-MAR-2002  
DEFINITION Drosophila melanogaster strain y; cn bw sp chromosome X clone

ACCESSION	AC007521	GI:19483792
VERSION	HTG, HTGS PHASE1.	
KEYWORDS	Drosophila melanogaster (fruit fly)	
SOURCE	Drosophila melanogaster	
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Spheroidea; Drosophilidae; Drosophila.	
REFERENCE	1 (bases 1 to 180919)	
AUTHORS	Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Lee,B., Lewis,S., Li,P., Lomoran,M.A., Mazda,P., Moshrefi,A.R., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Richards,S., Sethi,H., Swirskas,R.R., Wan,K.H., Webster,D., Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.	
TITLE	Sequencing of Drosophila melanogaster	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 180919)	
AUTHORS	Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomoran,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E., Swirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M.	
TITLE	Direct Submission	
JOURNAL	Submitted (10-MAY-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA	
REFERENCE	3 (bases 1 to 180919)	
AUTHORS	Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Lee,B., Lewis,S., Li,P., Lomoran,M.A., Mazda,P., Moshrefi,A.R., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E., Swirskas,R.R., Wan,K.H., Webster,D., Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M.	
TITLE	Direct Submission	
JOURNAL	Submitted (15-MAR-2002) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA	
COMMENT	On Mar 15, 2002 this sequence version replaced gi:7191020. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site ( <a href="http://www.fruitfly.org/sequence/">http://www.fruitfly.org/sequence/</a> ) or send email to <a href="mailto:bdg@fruitfly.berkeley.edu">bdg@fruitfly.berkeley.edu</a> . All contigs in this submission meet the following cutoffs: length >= 200 bases. * NOTE: This is a 'working draft' sequence. It currently consists of 61 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.	
	718	717: contig of 717 bp in length
	798	797: gap of unknown length
	1346	1345: contig of 548 bp in length
	1426	1425: gap of unknown length
	2042	2041: contig of 616 bp in length
	2122	2121: gap of unknown length
	2686	2685: contig of 564 bp in length
	2766	2765: gap of unknown length
	3509	3508: contig of 743 bp in length
	3589	3588: gap of unknown length
	4724	4723: contig of 1135 bp in length
	4804	4803: gap of unknown length
		5781: contig of 978 bp in length

BACR49A04 (D698) RPCI-98 49.A.4 map 10A2-10B2, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 61 unordered pieces.

```
* 5782 5861: gap of unknown length
* 5862 7147: contig of 1286 bp in length
* 7148 7227: gap of unknown length
* 7228 9276: contig of 2049 bp in length
* 9277 9356: gap of unknown length
* 9357 10426: contig of 1070 bp in length
* 10427 10506: gap of unknown length
* 10507 11545: contig of 1039 bp in length
* 11546 11625: gap of unknown length
* 11626 12578: contig of 953 bp in length
* 12579 12658: gap of unknown length
* 12659 13814: contig of 1156 bp in length
* 13815 13894: gap of unknown length
* 13895 15132: contig of 1238 bp in length
* 15133 15212: gap of unknown length
* 15213 16663: contig of 1451 bp in length
* 16664 16743: gap of unknown length
* 16744 17742: contig of 999 bp in length
* 17743 17822: gap of unknown length
* 17823 19718: contig of 1896 bp in length
* 19719 19798: gap of unknown length
* 19799 21740: contig of 1942 bp in length
* 21741 21820: gap of unknown length
* 21821 24129: contig of 2309 bp in length
* 24130 24209: gap of unknown length
* 24210 26333: contig of 2124 bp in length
* 26334 26413: gap of unknown length
* 26414 27844: contig of 1430 bp in length
* 27844 27923: gap of unknown length
* 27924 29790: contig of 1867 bp in length
* 29791 29870: gap of unknown length
* 29871 33117: contig of 3247 bp in length
* 33118 33197: gap of unknown length
* 33198 36063: contig of 2866 bp in length
* 36064 36143: gap of unknown length
* 36144 39335: contig of 3192 bp in length
* 39336 39415: gap of unknown length
* 39416 43414: contig of 3999 bp in length
* 43415 43494: gap of unknown length
* 43495 46045: contig of 2551 bp in length
* 46046 46125: gap of unknown length
* 46126 50185: contig of 4060 bp in length
* 50186 50265: gap of unknown length
* 50266 55215: contig of 4950 bp in length
* 55216 55295: gap of unknown length
* 55296 59535: contig of 4240 bp in length
* 59536 59615: gap of unknown length
* 59616 65396: contig of 5783 bp in length
* 65399 65478: gap of unknown length
* 65479 73776: contig of 8298 bp in length
* 73777 73856: gap of unknown length
* 73857 78055: contig of 4199 bp in length
* 78056 78135: gap of unknown length
* 78136 90995: contig of 12860 bp in length
* 90996 91075: gap of unknown length
* 91076 102049: contig of 10974 bp in length
* 102050 102129: gap of unknown length
* 102130 111675: contig of 9546 bp in length
* 111676 111756: gap of unknown length
* 111756 122893: contig of 11138 bp in length
* 122894 122973: gap of unknown length
* 122974 139619: contig of 16646 bp in length
* 139620 139699: gap of unknown length
* 139700 166353: contig of 26654 bp in length
* 166354 166433: gap of unknown length
* 166434 166864: contig of 431 bp in length
* 166865 166944: gap of unknown length
* 166945 167735: contig of 791 bp in length
* 167736 167815: gap of unknown length
* 167816 168509: contig of 694 bp in length
* 168510 168589: gap of unknown length
* 168590 168928: contig of 339 bp in length
* 168929 169008: gap of unknown length

* 169009 169748: contig of 740 bp in length
* 169749 169828: gap of unknown length
* 169829 170148: contig of 320 bp in length
* 170149 170228: gap of unknown length
* 170229 170610: contig of 382 bp in length
* 170611 170890: gap of unknown length
* 170891 170927: contig of 237 bp in length
* 170928 171007: gap of unknown length
* 171008 171769: contig of 762 bp in length
* 171770 171849: gap of unknown length
* 171850 172190: contig of 341 bp in length
* 172191 172270: gap of unknown length
* 172271 172592: contig of 322 bp in length
* 172593 172672: gap of unknown length
* 172673 173139: contig of 467 bp in length
* 173140 173219: gap of unknown length
* 173220 175301: contig of 2082 bp in length
* 175302 175381: gap of unknown length
* 175382 175850: contig of 469 bp in length
* 175851 175930: gap of unknown length
* 175931 176376: contig of 446 bp in length
* 176377 176456: gap of unknown length
* 176457 176939: contig of 483 bp in length
* 176940 177019: gap of unknown length
* 177020 177549: contig of 530 bp in length
* 177550 177629: gap of unknown length
* 177630 178146: contig of 517 bp in length
* 178147 178226: gap of unknown length
* 178227 178860: contig of 634 bp in length
* 178861 178940: gap of unknown length
* 178941 179549: contig of 609 bp in length
* 179550 179629: gap of unknown length
* 179630 180203: contig of 574 bp in length
* 180204 180283: gap of unknown length
* 180284 180919: contig of 636 bp in length.

FEATURES
      source
      1..180919
      /organism="Drosophila melanogaster"
      /mol_type="genomic DNA"
      /strain="Y: cn bw sp"
      /db_xref="taxon:7227"
      /chromosome="X"
      /map="10A2-10B2"
      /clone="BACR49A04 (D698) RPCI-98 49.A.4"
      /clone_lib="RPCI-98 (Roswell Park Cancer Institute
      Drosophila melanogaster BAC library, partial EcoRI in
      pBACE3.6"
      718..757
      /estimated_length=unknown
      1346..1425
      /estimated_length=unknown
      2042..2121

Query Match      31.4%; Score 967.8; DB 14; Length 180919;
Best Local Similarity 87.0%; Pred. No. Se-175;
Matches 1055; Conservative 0; Mismatches 82; Indels 76; Gaps 1;

QY 1697 CAAATACACACCCAAAGCGTTGGTGAACCAATTTTCGAGGCGCGAATGGCGACGTGCTT 1756
      |||
Db 44625 CAGATACACACCCAAAGCGTTGGTGAACCAATTTTCGAGGCGCGAATGGCGACGTGCTT 44566

QY 1757 CGCTACGGCCAGACGGGATCGGGCAAAACGCACACCATGGCGCGTGAAGTTTAATGGAAA 1816
      |||
Db 44565 CGCTACGGCCAGACGGGATCGGGCAAAACGCACACCATGGCGCGTGAAGTTTAATGGAAA 44506

QY 1817 GGTCCAGACTGCAAGAACGGCATCTACGCCATCGCGGCCAAGGATGCTTTGTGACCCCT 1876
      |||
Db 44505 GGTCCAGACTGCAAGAACGGCATCTACGCCATCGCGGCCAAGGATGCTTTGTGACCCCT 44446

QY 1877 GAATATCCCGGTTTACC CGCCCATGAATCTAGTCTCTCGGCCAGTTTCTTTGAGATTTA 1936
      |||
Db 44445 GAATATCCCGGTTTACC CGCCCATGAATCTAGTCTCTCGGCCAGTTTCTTTGAGATTTA 44386
```







1650 CGCTTCGACTACGCTTCAACGACACGCTGCGCAATGCGATGATATACAAATACACAGCC 1709  
853 CGTTTCGACTACACGCTTGACGAGAGTGCTCAATGCGTGTCTACGATCACAATGCT 912  
1710 AAGCGGTGGTCAAAACCAATTTTCAGGGCGGAATGGGACGCTGCTTCCGCTTACGGCCAG 1769  
913 CGTCCGTTGATCAGAACCAATGTTTCAGGGCGGCAATGCCACTGTTGTTCCGCTTACGGCAA 972  
1770 ACGGGATCGGGCAAAACGACACCAATGCGCGTGTGATTTAATGGAAGAGTGCAGGACTGC 1829  
973 ACTGCACGGGAAACGACCAATGCGCGGGAATTTCTCGAAGAGTTCAGGATGTC 1032  
1830 AAGAACCGCATCTACGCCATGCGCGGCAAGGATGCTTTGTGACCCCTGAATATGCGCGGT 1889  
1033 GGTACCGGGATCTACGCCATGCGAGCTCGCGATGCTTCGAGGAGTATCGCGCCGAG 1092  
1890 TACCGCGGCATGATCTAGTGTCTCGGCGAGTTCCTTTGAGATTTACAGTGGCAAGTGC 1949  
1093 TACCGGCAAAATGGGTGCCAAGATTACGTGACGCTTCTTCGAAATCTATGGCACCACAGGTG 1152  
1950 TTGATC---TTCTGTGCGACAAGCAGAACTGCGCGCTCTGGAGGATGTTAAACAGCAA 2006  
1153 TTGATCTCTGCTACCAACCAAGCCCATGCTGCGGGTCTTAGAGGATGCCAGGACGAG 1212  
2007 GTGCAAGTGTGGGACTCACCGAGAAGTGTGCTGATGCGCTCGAGGAGGTACTGAAGCTC 2066  
1213 GTGCTGTGTGGGCTTAACGAGATGCGGTGACCAAGTGGAGGATGCTCTGAGACTG 1272  
2067 ATCAGCAGGGAATGCTCGCGCAACATCCGCGCAGAGCTGGCGCAATCTCAATTCGTGTC 2126  
1273 ATTGAGCAGCGCAGCAAAAGAGCGCACTTCGCGCCAAACATCGGCGAAACGCGCAAGTCA 1332  
2127 GTTTCGACGCGCTTTTCCAGATTGCTGCTGCGCGCAGGCTCGACGAAGATCCATGGC 2186  
2333 GTTTCGACGCGCTTTTCAATAGCACTCCATTTTCCGANTCTCGGGGCCACACGGC 1392  
2187 AAGTCTCTGTTTCATGATCTGCGCGGCAATGAGCGGGCGTGGACACTTCCTCGCGCGAT 2246  
1393 AAGTCTCTTGTGACTTGGCGGCAATGAACGCGGGCGGATACGCAATCCGCGAT 1452  
2247 CGGCAAGCGGTATGAGGGTGGCGAGATTAAACAAATCGCTGCTGGCCCTCAAGGAGTGC 2306  
1453 GGTCAAACTCGCATGCGAGGAGCGGAGATCAATAAATCTCTGCTGGCCCTCAAGGAGTGC 1512  
2307 ATTGCTGTTGGCAACAGTTCGCGCACTTCGCTTCCGCTGCTCCCAACTCACCCAG 2366  
1513 ATTCAGCCCTCAGCGCGAGTGCAGTCACTTCTTCCGCTGCTCCCAAGTTCACCCAA 1572  
2367 GTGCTGCGGACTCGTTTCATTTGGCG---AGAAGAGCAAGAGTGCATATAGCCATGATC 2423  
1573 GTGCTGCGGACTCTTGTGCGCGCAAGAAAGAACAGACTCTGATGATTTGCCATGATA 1632  
2424 TCGCGGAGCTTAGTCTCTGCGAGCAGCTCAACAGCTCAACAGCTGCGCTATGCGGATGCTGTC 2483  
1633 TCGGCATCATGAGTGTGGTGGAGAAATACGCTCAACACTCTACGCTTACGCGACAGGGTT 1692  
2484 AAGGAGCTGGTTCAGGATATCG 2508  
1693 AAGGAGCTCATAGCAAGGAAGACG 1717

RESULT 13  
CQ577532  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CQ577532  
Sequence 5290 from Patent WO0171042.  
CQ577532  
CQ577532.1 GI:41640334  
Drosophila sp.  
Drosophila sp.  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

linear PAT 02-FEB-2004

REFERENCE 1  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
ORIGIN

Venter, J. C., Adams, M., Li, P. W. and Myers, E. W.  
Detection kits, such as nucleic acid arrays, for detecting the  
expression of 10,000 or more Drosophila genes and uses thereof  
Patent: WO 0171042-A 5290 27-SEP-2001;  
PE Corporation (NY) (US)  
Location/Qualifiers  
1. 4190  
/organism="Drosophila sp."  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:7242"

Query Match 18.3%; Score 564.6; DB 6; Length 4190;  
Best Local Similarity 68.1%; Pred. No. 9.8e-98;  
Matches 834; Conservative 0; Mismatches 379; Indels 12; Gaps 3;

Qy 1296 GCATTGAAAGAGTGGAGCGACTGAAGAGAAATCGCAGAGAGCGCGCCGACAGGCC 1355  
Db 1493 GTAGTGAAGGAGGTGAATCGCATGAAGAGCAGAGGAGGAGCGAGGGCTCGCCAGCG 1552  
Qy 1356 GAGATGAAGGAGGAGAGGTGGCGCTGATGAA CAGGATCCGGCAATCCAACTGGGAG 1415  
Db 1553 GAA CAGCTCCAGGAGAGGATGCATCGTCGCAATAATCCGGGGAATCCCACTGGGAG 1612  
Qy 1416 ACGCGCAAAATGATACCGGAATATCAGAGCAGCGTGAATTTGTGCCGCT-----GCTC 1469  
Db 1613 GTGTCGCTGATGCTGCCCAATACCGCTCCACCTTGATCTTTCTCCACTTCGATGCCGTG 1672  
Qy 1470 GATGCCAGGCGCTCGATGACCATCAGATCAGATGTCGTCGCAAGCGTCCCAITAGC 1529  
Db 1673 GATCCCAATGGAGCACTGTTACGAAATTTACGGTGTGTGCGAAGAACGACCATGAT 1732  
Qy 1530 CGAAGAGGTCAATCCAAAGAGATCGATGTCATTTCCGTCGCGCAAGGACATGCTC 1589  
Db 1733 CGCAAGGAGGAGAACTCCAAGAACCTGAGATCATCAGTTTCCAGTTCGCGACAGCGCTG 1792  
Qy 1590 ATGCTGACAGGCGCGCAGCAAGTGCACCTCAACAAAGTTTCTGGAGAGAACCAAGTTT 1649  
Db 1793 ATGCTCATGATGTCGCCCTCAAGGTGATCTCACCAGTTCTTGGAGACCACAATTC 1852  
Qy 1650 CGCTTCGACTACGCTTCAACGACACATGCGACAAATGCCATGCTATGATATCAACAGCC 1709  
Db 1853 CGTTTCGACTACAGTTTCGACGAGGAGTGTCCAAATGCGCTGCTACGATCAGCTGCT 1912  
Qy 1710 AAGCGCTTGTGAAAACCATTTTCGAGGCGGGAATGCGGACGCTGCTTCCGCTACGGCCAG 1769  
Db 1913 CGTCCGTTGATCAGAACCATGTTTCGAGGCGGCAATGCCACTTGTTCGCTTACGAGCAA 1972  
Qy 1770 ACGGATCGCGCAAAACGACACCATGCGCGGTGAGTTTAAATGGAAGGTGCAGGACTGC 1829  
Db 1973 ACTGGACGCAAAACGACACCATGCGCGGAGAAATTCCTTCGAAAGTTTCAAGATTGC 2032  
Qy 1830 AAGAACCGCATCTACGCGCATGCGGCGCAAGGATGCTTTTGTGACCCCTGAATATATCCGCGT 1889  
Db 2033 GGTACCGGGATCTACGCCATGCGAGCTCGCGATGCTCTTCGAGGAGGTATCGCGCCGAG 2092  
Qy 1890 TACCGCGCATGAAATCTAGTGTCTCGGCCAGTTTCTTTTGTGAGATTTACAGTGGCAAGTGC 1949  
Db 2093 TACCGGCAAAATGGGTGCCAAGATTACGTGCAAGTCTTCTTCGAAATCTATGGCACCAGGTG 2152  
Qy 1950 TTGATC---TTCTGTGCGCAAGCAGAACTCGCGCTCTCGGAGGATGTTAAACAGCAA 2006  
Db 2153 TTGATCTCTTGTACCAACAAAGCCCATGCTCGGGTCTTAGAGGATGCCAGGACAG 2212  
Qy 2007 GTGCAAGTGTGGGACTCACCGAGAAGGTGCTGATGCGCGTTCGAGGAGGTACTGAAGCTC 2066  
Db 2213 GTGCTGTGTGGGCTTAAACGAGATGCGCGTGAACCAAGTGGAGGATGCTCTGAGACTG 2272  
Qy 2067 ATCAGCAGCGCAATGCTGCCCGAAACATCCGCGCAGAGCTCGGCCAACTCCAATTCGTCG 2126  
Db 2273 ATTGAGCAGCGCAGCAAAAGAGCGCACTTCGCGCCAAACATCGCGCAACGCGCAAGTCTATCG 2332

```
QY 2127 CGTTCGACGCGCTTTTCAGATTGTGTCGGCCGCGAGGCTCGACGAGATCCATGCGC 2186
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2187 AAGTTCCTCGTTTCATCATCTGGCGGCAATGAGCGGGGTGACACTTCTCTCGGCCGAT 2246
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2247 CGGCAGACGCGTATGAGGGTGCAGAGATTAACTAATCTGCTGGCCCTCAAGGAGTGC 2306
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2307 ATTCGTGCGTTGGGCAACAGTCGCGCCCACTTCCCTCGTGTCTCCAAATCAACCCAG 2366
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2367 GTGCTCGCGACTCGTTTCATTGTCG--AGAAGACAGACGTCGATGATAGCATGATC 2423
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2424 TCGCCGGGACTTAGTCTCTCGGAGCACGCTCAACACGCTCGCTTATCGGATCGTGTTC 2483
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2484 AAGGAGCTCGTGTCAAGATATCG 2508
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2693 AAGGAGCTCATAGCCCAAGGACG 2717
```

RESULT 14  
AC006245/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## COMMENT

## JOURNAL

## COMMENT

## JOURNAL

## COMMENT

## JOURNAL

## COMMENT

## JOURNAL

## COMMENT

## JOURNAL

## COMMENT

## JOURNAL

## COMMENT

## JOURNAL

## COMMENT

## JOURNAL

## COMMENT

the following cutoffs: length >= 200 bases. P1 library location: 82-55.

\* NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. The true order of the pieces

\* is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence

\* as soon as it is available and the accession number will be preserved.

\* 1 588: contig of 588 bp in length

\* 589 668: gap of unknown length

\* 669 1229: contig of 561 bp in length

\* 1230 1309: gap of unknown length

\* 1310 82269: contig of 80960 bp in length

\* 82270 82349: gap of unknown length

\* 82350 82884: contig of 535 bp in length

\* 82885 82964: gap of unknown length

\* 82965 83532: contig of 568 bp in length

\* 83533 83612: gap of unknown length

\* 83613 84419: contig of 807 bp in length.

\* Location/Qualifiers

1. 84419

/organism="Drosophila melanogaster"

/mol\_type="genomic DNA"

/strain="y; cn bw sp"

/db\_xref="taxon:7227"

/chromosome="2"

/map="59D1-59D2"

/clone="DS07831 (D467)"

/clone\_lib="P1 library, partial Sau3A in pAd108acBII"

589. .568

/estimated\_length=unknown

1230. 1309

/estimated\_length=unknown

82270. .82349

/estimated\_length=unknown

82885. .82964

/estimated\_length=unknown

83533. .83612

/estimated\_length=unknown

ORIGIN

Query Match 18.3%; Score 564.6; DB 14; Length 84419;

Best Local Similarity 68.1%; Pred. No. 1.1e-97;

Matches 834; Conservative 0; Mismatches 379; Indels 12; Gaps 3;

QY 1296 GCATTGAAGAGGTGGAGCGACTGAAGAGAGATCCGAGAACGCGCGCAGCGCC 1355

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1356 GAGATGAAGGAGGAGGAGGTGGCGCTGATGAACAGGATCCGGGCAATCCAACTGGGAG 1415

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1416 ACGCGGCAATGATACGCAATATCAGAGCAGCTGGAATTTGTGCCCT-----GCTC 1469

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1470 GATGCCAGGCGCTCGATGACATCAGATCAGTGTGCGGCAAGCGTCCCAATTAGC 1529

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1530 CGCAGGAGGTCAATCCGAGGAGTGCATGTCATTTTCGTCGCCGCAAGCATGCTC 1589

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1590 ATCTGTCAGGCGCGCAGCAGGTGCACCTCACCAGTTCTTGGAGAACACCAAGTTT 1649

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1650 CGCTTCGACTACGCTTCAACGACGTCGCAATGCCATGGTATACAAATACACAGCC 1709

Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||



Search completed: May 12, 2006, 21:01:43  
Job time : 15365 secs

QY 1470 GATGGCCAGCCCGTCGATGACCATCAGATCACAGTGTGGTGGCGAAGCGTCCCATAGC 1529  
Db 133503 GATCCCAATGGAGGCACTGTTCAGCAAAATTACGGTGTGTGTGCGAAAACGACCCATGAGT 133444  
QY 1530 CGCAAGGAGGTCAATCGCAAGGAGATCGATGTCTATTTCCGTGCGCGCAAGGAGCATGCTC 1589  
Db 133443 CGCAAGGAGGAGAACTCCAGAAGACCTGGACATCATCACAGTTCCTGCGGACAGCCCTG 133384  
QY 1590 ATCGTGCAGAGCGCGCAGCAAGGTGCACTCACCAGTTCTCTGGAGAACCAACAAGTTT 1649  
Db 133383 ATCGTGCATGAGTTGGCGCTCAAGTGGATCTCACCAAGTTCTCGGAGCACCACAAATTC 133324  
QY 1650 CGCTTCGACTACGCTTCAACGACACGTGGGCAATGCCATGCTATACAAATACACAGCC 1709  
Db 133323 CGTTTCGACTACAGTTGCAACGAGGATGCTCCAATGCGCTGCTACGATCACACTGCT 133264  
QY 1710 AAGCCGTGTGTGAAAACCAATTTTCGAGGGCGGAATGGCGACGTGCTTCGCCTACGGCCAG 1769  
Db 133263 CGTCCGTTGATCAGAACCATTGTCGAGGGCGGCAATGCCACTTGTTCGCTTACGGACAA 133204  
QY 1770 ACGGATCGGGCAAAAACGACACCATGGGCGGTGAGTTTAAATGGAAGGTGCAGGACTGC 1829  
Db 133203 ACTGGCAGCGGAAAAACGACACCATGGGCGGAGAAATCTTCGGAAGGTTCAGGATTGC 133144  
QY 1830 AAGAAGCGCATCTACGCCATGGCGGCAAGGATGCTTTGTGACCTCGAATATGCCGCGT 1889  
Db 133143 GGTACCGGATCTACGCCATGGCAGCTCGCGATGCTTCGAGGAGGTATCGGCGCCGGAG 133084  
QY 1890 TACCGGCGCATGAATCTAGTCTGCTCGGCCAGTTCTTTGAGATTTACAGTGGCAGGTC 1949  
Db 133083 TACCGGCAATGGTCCGAGATTACGTGAGCTTCTTCGAAATCTATGGCACCAAGGTG 133024  
QY 1950 TTGATC---TTCTGTCCGACAGCAAGAACTGCGGCTCTCGAGGATGTTAAACAGCAA 2006  
Db 133023 TTGATCTCTGTCTACCAACCAAGCCATGCTCGGGTCTTAGAGGATGCCAGGCGAGCAG 132964  
QY 2007 GTGAGGTGGTGGACTACCGGAAGGTGCTGATGGCGTCGAGGAGGTACTGAAGCTC 2066  
Db 132963 GTCGTGGTGGTGGGCTAAACGGAGATGCGGTCGACCAAGTGGAGGATGCTCTGAGACTG 132904  
QY 2067 ATCCAGCAGCGCAATGCTGCCGCAATCCGCGACATCGCGGAGCTCGGCCAACTCCAAATTCGTCG 2126  
Db 132903 ATTGAGCAGCGCAGCAAGAGCGCACTTCGGCCAAACATCGCGGAAACGCCCAAGTCATCG 132844  
QY 2127 CGTTCGACGCGGTTTTCAGATTTGCTGCGGCGCAGGGCTCGACGAAGATCCATGGC 2186  
Db 132843 CGTTCCACGCGCTCTTTCAAATAGCACTCCACTTTCCCGATTCCTGGGGCCCAACAGGC 132784  
QY 2187 AAGTTCTGTTTATCGATCTGGGGGCAATGAGCGGGGCGTGGACACTTCCTCGGCGGAT 2246  
Db 132783 AAGTGTCTCTTTGTGACTTGGCGGGCAATGAACGCGGGCGGATACGCAATCCGCGCGAT 132724  
QY 2247 CGGCGAGCGGTATGAGGGTCCGAGATTAACAATCGCTGCTGGCCCTCAAGGAGTGC 2306  
Db 132723 CGTCAAACTCGCATCGAGGAGCCGAGATCAATAAATCTCTGTGGCCCTCAAGGAGTGC 132664  
QY 2307 ATTCTGCTGTGGCAAAACAGTCGGGCCACTTGGCCCTTCCTGTGTCTCCAAACTCACCCAG 2366  
Db 132663 ATTGAGGCTCAGCGCGGAGTCGAGTCACTTCCCTTCCTGGCTCCAAAGTTGACCCAA 132604  
QY 2367 GTGCTGCGGACTCGTTTCAATGGCG---AGAAGAGCAAGCGTGATGATAGCCATGATC 2423  
Db 132603 GTGCTGCGGACTCTCTTTGTGCGGCGCAAGAGAACAGACCTGATGATGCTCATGATA 132544  
QY 2424 TCGCGGGGACTTAGCTCTCGGAGCACAGCTCAACAGCTGCGCTATGCGGATCGTGTGC 2483  
Db 132543 TCGCCATCCATGAGCTGCGTGAGGAATACGCTCAACACTCTACGTTACGACAGAGGTT 132484  
QY 2484 AAGGAGCTCGGTCAAGGATATCG 2508  
Db 132483 AAGGAGCTCATAGCCAAAGGAGACG 132459

***This Page Blank (uspto)***

---

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model  
Run on: May 12, 2006, 16:33:29 ; Search time 1766 Seconds  
(without alignments)  
11627.360 Million cell updates/sec

Title: US-10-840-060-142  
Perfect score: 3081  
Sequence: 1 aaacataaaattgtgtgc.....gccgaatggcaagcgttagt 3081

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 21.\*  
1: geneseqn1980s.\*  
2: geneseqn1990s.\*  
3: geneseqn2000s.\*  
4: geneseqn2001as.\*  
5: geneseqn2001bs.\*  
6: geneseqn2002as.\*  
7: geneseqn2002bs.\*  
8: geneseqn2003as.\*  
9: geneseqn2003bs.\*  
10: geneseqn2003cs.\*  
11: geneseqn2003ds.\*  
12: geneseqn2004as.\*  
13: geneseqn2004bs.\*  
14: geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3081	100.0	3081	4	ABL02575 Drosophil
2	3081	100.0	3081	10	ADK11353 Drosophil
3	1547.8	50.2	8043	4	ABL02574 Drosophil
4	564.6	18.3	2190	4	ABL05367 Drosophil
5	564.6	18.3	4190	4	ABL05366 Drosophil
6	505.8	16.4	1970	4	ABL05833 Drosophil
7	505.8	16.4	3370	4	ABL05832 Drosophil
8	454.6	14.8	1292	6	AAD24080 Human mit
9	454.6	14.8	1292	6	AAD24080 Human mit
10	454.6	14.8	1292	12	AD052609 Human mit
11	454.6	14.8	1421	6	AAD24083 Human mit
12	454.6	14.8	1421	10	ADE10088 DNA encod
13	454.6	14.8	1421	12	AD052615 Human mit
14	454.6	14.8	2172	6	AAD24086 Human mit
15	454.6	14.8	2172	10	ADE10094 DNA encod
16	454.6	14.8	2172	12	AD052621 Human mit
17	453	14.7	2865	6	AAS94953 Human DNA
18	449.8	14.6	1175	6	AAD24081 Human mit
19	449.8	14.6	1175	10	ADE10084 DNA encod

20	449.8	14.6	1175	12	ADO52611 Human mit
21	449.8	14.6	1304	6	AAD24084 Human mit
22	449.8	14.6	1304	10	ADE10090 DNA encod
23	449.8	14.6	1304	12	ADO52617 Human mit
24	449.8	14.6	2740	6	ABL65400 Lung canc
25	449.8	14.6	2740	6	ABL65848 Lung canc
26	449.8	14.6	2740	6	ABL64767 Lung canc
27	449.8	14.6	2740	6	ABL70308 Human lun
28	449.8	14.6	2740	6	ABK70303 Human mit
29	449.8	14.6	2740	6	AAD24087 Human mit
30	449.8	14.6	2740	9	ACA90156 Human cdn
31	449.8	14.6	2740	10	ADE10096 DNA encod
32	449.8	14.6	2740	10	AAD54025 Human col
33	449.8	14.6	2740	12	ADO52623 Human gen
34	449.8	14.6	2740	13	ADT90126 Human gen
35	449.8	14.6	2740	13	ADT90283 Human gen
36	449.8	14.6	2740	13	ADT89995 Human gen
37	449.8	14.6	2740	13	ADT89892 Human gen
38	449.8	14.6	2825	12	ADQ09301 Human KNS
39	449.8	14.6	2825	13	ADR25467 Breast ca
40	449.8	14.6	2825	14	ADX06781 Cyclin-de
41	449.8	14.6	2825	14	ADY16204 DNA encod
42	445.2	14.4	2892	12	ADQ22362 Human eof
43	441	14.3	2857	12	ADQ85075 Human tum
44	441	14.3	2857	13	ACN37373 Tumour-as
45	436.2	14.2	2741	3	AAF15699 Human pro

ALIGNMENTS

RESULT 1  
ABL02575  
ID ABL02575 standard; cdna; 3081 BP.  
XX  
AC ABL02575;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 2207.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PE (PEKE ) PE CORP NY.  
XX  
Venter JC, Adams M, Li PWD, Myers EW;  
XX  
WPI: 2001-656860/75.  
XX  
P-PSDB; ABB58472.  
XX  
New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.  
XX  
Claim 1; SEQ ID NO 2207; 21pp + Sequence Listing; English.  
XX  
The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention



CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA sequences (AB101840-AB116175) and the encoded proteins (AB157737-AB172072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences									
XX	Sequence	3081 BP;	803 A;	817 C;	868 G;	593 T;	0 U;	0 Other;	
Query Match									
Best Local Similarity 100.0%; Score 3081; DB 4; Length 3081;									
Matches 3081; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	AAACTAAAAAATTTGTTGCTGACATCTGGTGCCTTGCAGAACTATTTCTAGCAGATTTT	60						
DB	1	AAACTAAAAAATTTGTTGCTGACATCTGGTGCCTTGCAGAACTATTTCTAGCAGATTTT	60						
QY	61	GTGATATTTTCGTTGATCGGTGCGATTAATCCGCCAGTTTTTTTTTAAATGGAAGTGCT	120						
DB	61	GTGATATTTTCGTTGATCGGTGCGATTAATCCGCCAGTTTTTTTTTAAATGGAAGTGCT	120						
QY	121	AAACACATTTGATCGGTTCGGAAGATAGCAGGAAGAGCCAGCGGCTGCGTTTTTCCTT	180						
DB	121	AAACACATTTGATCGGTTCGGAAGATAGCAGGAAGAGCCAGCGGCTGCGTTTTTCCTT	180						
QY	181	TTTGTATTCGTTGCCAGACGCAACGAAACGACATTTGGCATTTTGAATTCAGCACAAC	240						
DB	181	TTTGTATTCGTTGCCAGACGCAACGAAACGACATTTGGCATTTTGAATTCAGCACAAC	240						
QY	241	ACACATATTAACCGCCAGCCGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	300						
DB	241	ACACATATTAACCGCCAGCCGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC	300						
QY	301	AAAAAACAAGACGCTCTCGCGGACCTCGAAGCAGTTCGGTTTCGATTTAGTTTCAATGC	360						
DB	301	AAAAAACAAGACGCTCTCGCGGACCTCGAAGCAGTTCGGTTTCGATTTAGTTTCAATGC	360						
QY	361	CTTGAATTCGTTTCGGGCTTTAGTTTTCACAAAGTTTATCGTCTGTCGTAAGAACGAAA	420						
DB	361	CTTGAATTCGTTTCGGGCTTTAGTTTTCACAAAGTTTATCGTCTGTCGTAAGAACGAAA	420						
QY	421	TAAATATTTTCGACCTTAAATAATCTGACTAAATTTGTTTGTATTTATTTATTT	480						
DB	421	TAAATATTTTCGACCTTAAATAATCTGACTAAATTTGTTTGTATTTATTTATTT	480						
QY	481	TAGGCACATTTTGACACACAAACGCTAGTTACTATCTACGCTAACGGAACCTCTCT	540						
DB	481	TAGGCACATTTTGACACACAAACGCTAGTTACTATCTACGCTAACGGAACCTCTCT	540						
QY	541	GCAAGCAGTGAAGTTGCTGCTCCATCAAGCAGTACTCGGAGTTAAAGCAGGATAAGCCG	600						
DB	541	GCAAGCAGTGAAGTTGCTGCTCCATCAAGCAGTACTCGGAGTTAAAGCAGGATAAGCCG	600						
QY	601	GAGAAAGAGAGATCGGTGGAGATAGAGATATACAGTGGAGTCAAGAGAGAGGA	660						
DB	601	GAGAAAGAGAGATCGGTGGAGATAGAGATATACAGTGGAGTCAAGAGAGAGGA	660						
QY	661	TCATGGACATGATTTACGTTGGGCGAGCGCTCAAGATCAAGCGAGTTCGCGCTCC	720						
DB	661	TCATGGACATGATTTACGTTGGGCGAGCGCTCAAGATCAAGCGAGTTCGCGCTCC	720						
QY	721	ACATGGCCGTTGGTGGCGTATCAACAGTCCGGCAGTCAAGTCAAGTCAAGTCAAGT	780						
DB	721	ACATGGCCGTTGGTGGCGTATCAACAGTCCGGCAGTCAAGTCAAGTCAAGTCAAGT	780						
QY	781	AGCGGGCGAAACGAAAGGCGAGGAGTGAATCTGGAGCCATCTACGCTCAATCCGG	840						
DB	781	AGCGGGCGAAACGAAAGGCGAGGAGTGAATCTGGAGCCATCTACGCTCAATCCGG	840						
QY	841	AGCTAAATCAAGATCTGTCGAAACGACGCGCGCCCGGAGCCCAAGAAACAAGCCACCG	900						
DB	841	AGCTAAATCAAGATCTGTCGAAACGACGCGCGCCCGGAGCCCAAGAAACAAGCCACCG	900						
QY	901	CGCCGATGAACCTCTCGGTAATCCACAAATCGGCTATCGGTGGCAATCTCACCGCC	960						

DB	901	CGCCGATGAACCTCTCGGTAATCCACAAATCGGCTATCGGTGGCAATCTCACCGCC	960						
QY	961	GTATGACCATCGCGGAAACATGCTGAAACAGATCCAGGAAAGCAGTCTGATCTCCAAATC	1020						
DB	961	GTATGACCATCGCGGAAACATGCTGAAACAGATCCAGGAAAGCAGTCTGATCTCCAAATC	1020						
QY	1021	CGATTGTCAGCAGCAATAGCGTGAATCAAAACAGCAACTCCAACTACGCGCGCGGAG	1080						
DB	1021	CGATTGTCAGCAGCAATAGCGTGAATCAAAACAGCAACTCCAACTACGCGCGCGGAG	1080						
QY	1081	GTGTGGCACCACCAACGTCGACGACCACTGATTTACAGCGTCCACGGTACTTCGCAAGCTG	1140						
DB	1081	GTGTGGCACCACCAACGTCGACGACCACTGATTTACAGCGTCCACGGTACTTCGCAAGCTG	1140						
QY	1141	CTACCGGCGACGACGAGCAAGGATTCGCTTCGGCGGTGCTTAATAACATTTGCCCAATC	1200						
DB	1141	CTACCGGCGACGACGAGCAAGGATTCGCTTCGGCGGTGCTTAATAACATTTGCCCAATC	1200						
QY	1201	CCAGCGGCGACGAGTCTGCTCGCGGCGACAGGAGTCCGCACTGCGGCGCACACCC	1260						
DB	1201	CCAGCGGCGACGAGTCTGCTCGCGGCGACAGGAGTCCGCACTGCGGCGCACACCC	1260						
QY	1261	AGGAGCTGGCGGCGCTAGTACCCGCGGATCGCACGCAATTCGAAAGAGTGGAGCGACTGA	1320						
DB	1261	AGGAGCTGGCGGCGCTAGTACCCGCGGATCGCACGCAATTCGAAAGAGTGGAGCGACTGA	1320						
QY	1321	AGGAGATTCGCGAGAGCGCGCCGACAGGCGGAGATGAAGGAGAGAGTGGCGC	1380						
DB	1321	AGGAGATTCGCGAGAGCGCGCCGACAGGCGGAGATGAAGGAGAGAGTGGCGC	1380						
QY	1381	TGATGAACAGGATTCGCGGCAATCCAACTGGGAGAGCGCGCAATGATACGCAATATC	1440						
DB	1381	TGATGAACAGGATTCGCGGCAATCCAACTGGGAGAGCGCGCAATGATACGCAATATC	1440						
QY	1441	AGACACGCTGGAAATTTGTCGCTCGATGCGCGAGCGCTCGATGACCATCAGATCA	1500						
DB	1441	AGACACGCTGGAAATTTGTCGCTCGATGCGCGAGCGCTCGATGACCATCAGATCA	1500						
QY	1501	CAGTGTGCGTGGCGACGCTCCATTTAGCCGAGAGAGTCAATCGCAAGAGATCGATG	1560						
DB	1501	CAGTGTGCGTGGCGACGCTCCATTTAGCCGAGAGAGTCAATCGCAAGAGATCGATG	1560						
QY	1561	TCATTTGCGTGGCGCAAGCAGTCTCATCTGTCAGCAGCGCGCAGCAAGTCCAGC	1620						
DB	1561	TCATTTGCGTGGCGCAAGCAGTCTCATCTGTCAGCAGCGCGCAGCAAGTCCAGC	1620						
QY	1621	TCACCAAGTTCTGGAGAACCAAGTTTCGCTTCGACTACGCTTCAACGACGCTGCG	1680						
DB	1621	TCACCAAGTTCTGGAGAACCAAGTTTCGCTTCGACTACGCTTCAACGACGCTGCG	1680						
QY	1681	ACATGCCATGTTATCAAAATACACAGCAGCGCTGGTGGTGAACCATTTTCAGGGCG	1740						
DB	1681	ACATGCCATGTTATCAAAATACACAGCAGCGCTGGTGGTGAACCATTTTCAGGGCG	1740						
QY	1741	GAATGGCGAGTCTGCTCGCTTCAGCGCAGCGGATCGCGCAAAACGCAACCATGGCG	1800						
DB	1741	GAATGGCGAGTCTGCTCGCTTCAGCGCAGCGGATCGCGCAAAACGCAACCATGGCG	1800						
QY	1801	GTGAGTTTAAATGGAAGGTCAGGACTGCAAGAACGSCATCTAGCCCATGGCCCAAGG	1860						
DB	1801	GTGAGTTTAAATGGAAGGTCAGGACTGCAAGAACGSCATCTAGCCCATGGCCCAAGG	1860						
QY	1861	ATGCTTTTGTGACCTGAAATATGCGCGTTACCGCGCATGAATCTAGTCTCTCGGCCA	1920						
DB	1861	ATGCTTTTGTGACCTGAAATATGCGCGTTACCGCGCATGAATCTAGTCTCTCGGCCA	1920						
QY	1921	GTGTTCTTTGAGATTTACAGTGGCAAGTCTTTCGATCTTCTGTCGCAAGCAGAACTGC	1980						
DB	1921	GTGTTCTTTGAGATTTACAGTGGCAAGTCTTTCGATCTTCTGTCGCAAGCAGAACTGC	1980						
QY	1981	GCCTCTCTGGAGGTGTTAAACAGCAAGTGCAGTGGTGGGACTCACCGAAGGTGCTG	2040						

Db 1981 GCGTCTGGAGGATGGTAAACAGCAAGTCAGAGTGGTGGACTCACCGAAGGTTGGTCG 2040  
Qy 2041 ATGCGCTGAGAGGTTACTGAAGCTCATCAGACAGCAATGTGTCGCGCAACATCCGCGCC 2100  
Db 2041 ATGCGCTGAGAGGTTACTGAAGCTCATCAGACAGCAATGTGTCGCGCAACATCCGCGCC 2100  
Qy 2101 AGAGCTCGGCAACTCCAAATTCGTCGCTTCGACAGCGCGTTTCCAGATTGTGCTGGCG 2160  
Db 2101 AGAGCTCGGCAACTCCAAATTCGTCGCTTCGACAGCGCGTTTCCAGATTGTGCTGGCG 2160  
Qy 2161 CGCAGGCTCGACGAAGATCCATGGCAAGTTCTGTTTCATCGATCTGCGCGGCAATGAGC 2220  
Db 2161 CGCAGGCTCGACGAAGATCCATGGCAAGTTCTGTTTCATCGATCTGCGCGGCAATGAGC 2220  
Qy 2221 GGGCGGTGAGCACTTCTCGCGCGATCGGACAGCGGTATGAGAGGTGCGGAGATTAACA 2280  
Db 2221 GGGCGGTGAGCACTTCTCGCGCGATCGGACAGCGGTATGAGAGGTGCGGAGATTAACA 2280  
Qy 2281 AATCGCTGCTGGCCCTCAAGAGTGCATTCTGCTGGTGGGCAACAGTCGGGCCCACTTGC 2340  
Db 2281 AATCGCTGCTGGCCCTCAAGAGTGCATTCTGCTGGTGGGCAACAGTCGGGCCCACTTGC 2340  
Qy 2341 CCTTCCGTGTCTCCAACTCACCCAGGTGCTGCGCGACTCGTTTCATTGGCGGAAGAGCA 2400  
Db 2341 CCTTCCGTGTCTCCAACTCACCCAGGTGCTGCGCGACTCGTTTCATTGGCGGAAGAGCA 2400  
Qy 2401 AGAGTGCATGATAGCATGATCTGCGCGGGAATTAGTCTCTCGGACACACGCTCAACA 2460  
Db 2401 AGAGTGCATGATAGCATGATCTGCGCGGGAATTAGTCTCTCGGACACACGCTCAACA 2460  
Qy 2461 CGCTGGCTATGCGGATCGTGTCAAGAGCTGTGTCAAGGATATCTGTCGAAGTTTGGC 2520  
Db 2461 CGCTGGCTATGCGGATCGTGTCAAGAGCTGTGTCAAGGATATCTGTCGAAGTTTGGC 2520  
Qy 2521 CTGCGCGGACACCGAGCCCATCGAGATCACGACACGAGGAGGAGGAGGCTCAACA 2580  
Db 2521 CTGCGCGGACACCGAGCCCATCGAGATCACGACACGAGGAGGAGGAGGCTCAACA 2580  
Qy 2581 TGTGTCATCGCACTCGCATCAGTGCATCCCAATTCGATGACCGCGGCGAGCTCGA 2640  
Db 2581 TGTGTCATCGCACTCGCATCAGTGCATCCCAATTCGATGACCGCGGCGAGCTCGA 2640  
Qy 2641 ATAATCAGCTGTCTCGGCTCTCATCATCTCGGGGGGGTTCATTCACAAATTAATA 2700  
Db 2641 ATAATCAGCTGTCTCGGCTCTCATCATCTCGGGGGGGTTCATTCACAAATTAATA 2700  
Qy 2701 ACAACAAAGAACGAAACCGCGCAACATGACCTGGCCATGCTGAGTTGCTGAGCG 2760  
Db 2701 ACAACAAAGAACGAAACCGCGCAACATGACCTGGCCATGCTGAGTTGCTGAGCG 2760  
Qy 2761 AACACAGATGTCGACGAGTGAATGTGACGACACGAGGCCATTCGACGACCTGACGAGA 2820  
Db 2761 AACACAGATGTCGACGAGTGAATGTGACGACACGAGGCCATTCGACGACCTGACGAGA 2820  
Qy 2821 CGGAGAGATGTGTTGGAGTATCATCGACCGTTAATGCCACATCTGGAGACCTTCTCG 2880  
Db 2821 CGGAGAGATGTGTTGGAGTATCATCGACCGTTAATGCCACATCTGGAGACCTTCTCG 2880  
Qy 2881 CCGAGTCAAGGGCGGTGAATCTGACCAATATGTGAGTACGACGAGGAGTCTGCTACT 2940  
Db 2881 CCGAGTCAAGGGCGGTGAATCTGACCAATATGTGAGTACGACGAGGAGTCTGCTACT 2940  
Qy 2941 GCAACCGGGCGAGTGTCTTCGAGTGTGGACATCGCCATCCAGTCCGCGACA 3000  
Db 2941 GCAACCGGGCGAGTGTCTTCGAGTGTGGACATCGCCATCCAGTCCGCGACA 3000  
Qy 3001 TGAATGCCGATATCGGCGAAGTTGGCCAGAGAGATGCTGTGCTGAGCTTCAATT 3060  
Db 3001 TGAATGCCGATATCGGCGAAGTTGGCCAGAGAGATGCTGTGCTGAGCTTCAATT 3060  
Qy 3061 CGCCGAATGCGAGCGTTAGT 3081  
Db 3061 CGCCGAATGCGAGCGTTAGT 3081

## RESULT 2

ADK11353

ID ADK11353 standard; DNA; 3081 BP.

XX AC ADK11353;

XX XX

DT 06-MAY-2004 (first entry)

XX XX

Drosophila kinesin-like protein KIF2 homolog gene.

de; gene; cytostatic; cardiovascular; immunosuppressive; nephrotropic;  
antirheumatic; antiarthritic; dermatological; antiparasitic;  
antiinflammatory; fungicide; gene therapy; Drosophila; diagnosis;  
cardiovascular disorder; autoimmune disease; glomerulonephritis;  
rheumatoid arthritis; dermatological disorder; psoriasis;  
inflammatory disorder; malaria; emphysema; alopecia.

Drosophila melanogaster.

WO2003040301-A2.

15-MAY-2003.

23-OCT-2002; 2002WO-GB004780.

05-NOV-2001; 2001GB-00026506.

27-NOV-2001; 2001GB-00028384.

11-FEB-2002; 2002GB-00003185.

(CYCL-) CYCLACEL LTD.

Deak P, Frenz L, Glover D, Midgley C;

WPI; 2003-441540/41.

P-PSDB; ADK11354.

New Drosophila polypeptides and polynucleotides, useful for diagnosing,  
preventing and/or treating disorders, such as cancer, glomerulonephritis,  
rheumatoid arthritis, psoriasis, malaria, emphysema and alopecia.

Example 9; Page 129-130; 265pp; English.

The invention relates to novel Drosophila species DNA sequences and their  
encoded proteins with their corresponding human homologues. The proteins  
or their encoding polynucleotides are useful in a method of prevention,  
treatment or diagnosis of a disease in an individual, and used to  
identify a substance capable of binding to the polypeptide or modulating  
the function of the polypeptide comprising incubating the polypeptide  
with a candidate substance and determining whether the substance binds to  
the polypeptide. The compositions are administered to an individual in  
need of such treatment. The method of diagnosis, in which the presence or  
absence of a polynucleotide is detected in a biological sample, comprises  
brining the biological sample containing the nucleic acid such as DNA or  
RNA into contact with a probe comprising a fragment of at least 15  
nucleotides of the polynucleotide, and detecting any duplex formed  
between the probe and nucleic acid in the sample. The method also  
comprises providing an antibody capable of binding to the polypeptide,  
incubating a biological sample with the antibody to allow the formation  
of an antibody-antigen complex, and determining whether antibody-antigen  
complex comprising the antibody is formed. The disease comprises a  
proliferative disease such as cancer. The antibody or identified  
substance is also useful in inhibiting the function of a polypeptide  
and/or regulating a cell division cycle function. The diseases also  
include cardiovascular disorders, autoimmune diseases such as  
glomerulonephritis and rheumatoid arthritis, and dermatological disorders  
such as psoriasis, inflammatory, fungal, and parasitic disorders such as  
malaria, emphysema and alopecia. This sequence represents one of the  
Drosophila genes of the invention.

Sequence 3081 BP; 803 A; 817 C; 868 G; 593 T; 0 U; 0 Other;





Db 1959 ATCTACGATAACCGAATCTCTCTGCAAGCAGTGGAACTGTCTGTCCATCAAGCAGTAC 2018  
Qy 576 TCGGAGTTAAACCGAGTAAGCCGGGAGAAAGAGATCGGTGGAGATAGAGATA 635  
Db 2019 TCGGAGTTAAACCGAGTAAGCCGGGAGAAAGAGATCGGTGGAGATAGAGATA 2078  
Qy 636 TACAGGTGAGTCAAGAGAAAGGATCATGGACATGATTACGGTGGGAGAGCGTCAAG 695  
Db 2079 TACAGGTGAGTCAAGAGAAAGGATCATGGACATGATTACGGTGGGAGAGCGTCAAG 2138  
Qy 696 ATCAAGCGAGCGATGGCGCGCTCCACATGCGCGGTGGTGGCGGTGATCAACAGTCGGGC 755  
Db 2139 ATCAAGCGAGCGATGGCGCGCTCCACATGCGCGGTGGTGGCGGTGATCAACAGTCGGGC 2198  
Qy 756 AAGTGCATCAGAGTCAAGTGGTACGAGCGCGGGAACGAAAGGGCAAGAGGTAGAACTG 815  
Db 2199 AAGTGCATCAGAGTCAAGTGGTACGAGCGCGGGAACGAAAGGGCAAGAGGTAGAACTG 2258  
Qy 816 GACGCCATCTCAAGTCCGAGGCTAATGCAAGATCTGTGCAACAGCAGCCCGCC 875  
Db 2259 GACGCCATCTCAAGTCCGAGGCTAATGCAAGATCTGTGCAACAGCAGCCCGCC 2318  
Qy 876 CCGGAGCCCAAGAAACAGCCACCGCGCGATGAACCTCTCGGTAAATCCCAACAATCG 935  
Db 2319 CCGGAGCCCAAGAAACAGCCACCGCGCGATGAACCTCTCGGTAAATCCCAACAATCG 2378  
Qy 936 GCTATCGGTGCAATCTCACCAGCGGTATGACCATGGCGGGAACATCTGCAAGATC 995  
Db 2379 GCTATCGGTGCAATCTCACCAGCGGTATGACCATGGCGGGAACATCTGCAAGATC 2438  
Qy 996 CAGGAAAGCCAGTCCGATTCCTCAATCCGATTTGTGACAGCAATAGCGTGAATACAAACAGC 1055  
Db 2439 CAGGAAAGCCAGTCCGATTCCTCAATCCGATTTGTGACAGCAATAGCGTGAATACAAACAGC 2498  
Qy 1056 AACTCCAACTACCGCGCGGAGTGGTGGCACCAACAGTTCGACGACCACTGGATTA 1115  
Db 2499 AACTCCAACTACCGCGCGGAGTGGTGGCACCAACAGTTCGACGACCACTGGATTA 2558  
Qy 1116 CAGCGTCCAGCTACTCCGAGCTGTACCGGCGAGCAGCAGAGATGCTCGCGG 1175  
Db 2559 CAGCGTCCAGCTACTCCGAGCTGTACCGGCGAGCAGCAGAGATGCTCGCGG 2618  
Qy 1176 GTGCCATAAACAATTCCTCAATCCAGCGCGGAGTGTGTCGCGCGGCAAA 1235  
Db 2619 GTGCCATAAACAATTCCTCAATCCAGCGCGGAGTGTGTCGCGCGGCAAA 2678  
Qy 1236 GGAGTCCGCACTGCGGCCAACCAACCGAGGAGTGGCGCGGTAGTACCCGCGGATCGCAC 1295  
Db 2679 GGAGTCCGCACTGCGGCCAACCAACCGAGGAGTGGCGCGGTAGTACCCGCGGATCGCAC 2738  
Qy 1296 GCATTGAAAGAGTGGAGGAGTGAAGAGATCGCGAGAGCGACGCGCGGACAGGCC 1355  
Db 2739 GCATTGAAAGAGTGGAGGAGTGAAGAGATCGCGAGAGCGACGCGCGGACAGGCC 2798  
Qy 1356 GAGATGAAAGAGGAGGAGTGGCGCTGATGAACAGAGATCCGGGCAATCCAACTGGGAG 1415  
Db 2799 GAGATGAAAGAGGAGGAGTGGCGCTGATGAACAGAGATCCGGGCAATCCAACTGGGAG 2858  
Qy 1416 ACGGCGCAATGATACGCGCAATATCAGAGCAGCTGGAAATTTGTGCCGCTCTCGATGGC 1475  
Db 2859 ACGGCGCAATGATACGCGCAATATCAGAGCAGCTGGAAATTTGTGCCGCTCTCGATGGC 2918  
Qy 1476 CAGGCGCTCGATGACCATCAGATCAGATGAGTGGCGTGGCAGAGCTCCCATAGCGGCAAG 1535  
Db 2919 CAGGCGCTCGATGACCATCAGATCAGATGAGTGGCGTGGCAGAGCTCCCATAGCGGCAAG 2978  
Qy 1536 GAGTCAATCGCAAGAGATCGATGCTCATTTTCGGTGGCGCGCAAGGACATGCTCATCGTG 1595  
Db 2979 GAGTCAATCGCAAGAGATCGATGCTCATTTTCGGTGGCGCGCAAGGACATGCTCATCGTG 3038  
Qy 1596 CACGAGCGCGCAGCAGGTCGACTCAACCAAGTTCTCTGGAGAAACCAAGTTTCGCTTC 1655

Db 3039 CACGAGCGCGCAGCAGGTCGACCTCACCAAGTTCTCTGGAGAAACCAAGTTTCGCTTC 3098  
Qy 1656 GACTACGCTTCAACGACACGTCGACCAATGCCATGGTATACAAATACACAGCCAGCC 1714  
Db 3099 GACTACGCTTCAACGACACGTCGACCAATGCCATGGTATACAAATGAGTATCAATCC 3157  
RESULT 4  
ABL05367  
ID ABL05367 standard; cDNA; 2190 BP.  
XX ABL05367;  
AC  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 10583.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
pharmaceutical; gene; ss.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US009231.  
XX  
PR 23-MAR-2000; 2000US-0191637P.  
PR 11-JUL-2000; 2000US-00614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI; 2001-656860/75.  
DR P-PSDB; ABB61264.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.  
PS Claim 1; SEQ ID NO 10583; 21bp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
SQ Sequence 2190 BP; 586 A; 574 C; 626 G; 404 T; 0 U; 0 Other;  
Query Match 18.3%; Score 564.6; DB 4; Length 2190;  
Best Local Similarity 68.1%; Pred. No. 3.1e-134;  
Matches 834; Conservative 0; Mismatches 379; Indels 12; Gaps 3;  
Qy 1296 GCATTGAAAGAGTGGAGCGCTGAAGAGAAATCGCAGAGAGCGACGCGCCGACAGGCC 1355  
Db 493 GTAGTAGGAGGTGAATCGCATGAGAGAGAGGAGGAGCGAAGGGCTCGCAGCG 552  
Qy 1356 GAGATGAAAGAGGAGAGTGGCGCTGATGAACAGAGATCCGGCAATCCAACTGGGAG 1415  
Db 553 GAACAGCTCCAGGAGAGGATGCACTGCGTGGCAATTAATCGGGGAATCCCACTGGGAG 612  
Qy 1416 ACGGCGCAATGATACCGCAATATCAGAGCAGCTGGAAATTTGTGCCGCT-----GCTC 1469  
Db 613 GTGTGGTGTGTGCTGCGCCAAATACCGCTCCACCTTGATCTTTTCTCCACTTCGATGCTG 672

QY 1470 GATGGCCAGCGCTGATGACCATCAGATCAGATGCGTGGCGAAGCGTCCCATTTAGC 1529  
| | | | |  
Db 673 GATCCCAATGGAGCACTGTTAGCAAAATTACGGTGTGTGTGCGAAACACCCATGATG 732  
| | | | |  
QY 1530 CGCAAGGAGGTCAATCGCAAGGAGATCGATGTCTCATTTTCGGTGGCGCGCAAGGACATGCTC 1589  
| | | | |  
Db 733 CGCAAGGAGGAGAACTCCAAGAACCTTGGACATCATCACAGTCCCAAGTGGCGACAGCCTG 792  
| | | | |  
QY 1590 ATCGTGACAGCGCGCGCAGCAAGGTGCGACCTCACCAAGTTCTCTGGAGAACCAAGATTT 1649  
| | | | |  
Db 793 ATCGTCCATGATTTGGCGCTCAAGGTGGATCTCACCAAGTTCTCTGGAGCACCACAAAATC 852  
| | | | |  
QY 1650 CGTTTCGACTACCGCTTCAACGACACGTGCGACCAATGCCCATGTATACAAATACACAGCC 1709  
| | | | |  
Db 853 CGTTTCGACTACACGTTCGACGAGGAGTGTCTCAATGCGCTGTCTACGATCATCATGCT 912  
| | | | |  
QY 1710 AAGCCGTGTGTAAGAACCATTTTCGAGGGCGGGAATGGCAGCTGCTTCGCTACGGCCAG 1769  
| | | | |  
Db 913 CGTCCGTTGATCAGAACCATGTTTCGAGGGCGGCAATGCCACTTGTTCGCTTACGGACAA 972  
| | | | |  
QY 1770 ACGGATCGGGCAAAACGCAACCATGGGCGGTGAGTTTAAATGGAAGGTGCAAGGACTGC 1829  
| | | | |  
Db 973 ACTGGCAGCGGAAAAACGACACCATGGGCGGAGAAATTCCTCGGAAAGGTTACAGGATTC 1032  
| | | | |  
QY 1830 AAGAACGGCATCTACGCCATGCGCGGCAAGGATGCTTTGTGACCCCTGATATGCCCGT 1889  
| | | | |  
Db 1033 GGTACCGGATCTACGCCATGGCAGCTCGCGATGCTTCGAGGAGTATCGGCCCGGAG 1092  
| | | | |  
QY 1890 TACCGGCCCATGAATCTAGTCGTCTCGGCCAGTTTCTTTGAGATTTTACAGTGCAGAGTTC 1949  
| | | | |  
Db 1093 TACCGGCAATGGTGCCAGATTAGTGCAGCTTCTTCGAATCTATGGCACCAGGTG 1152  
| | | | |  
QY 1950 TTGATC---TTCTGTCCGAAGAGCAAACTCGCGTCTCGGAGTGGTAAACAGCAA 2006  
| | | | |  
Db 1153 TTGATCTCTGTACTACCAACAGCCCATCTCGCGGTCTTAGAGGATCGCGGACGAG 1212  
| | | | |  
QY 2007 GTGCAGGTGGTGGATCTCCGAGAGAGTGGTTCATGGGTGCGAGAGGTACTGAAGTTC 2066  
| | | | |  
Db 1213 GTCGTGGTGGTGGCCCTAACCGAGATGCGCGGTGACCAAGTGGAGATGCTCTGAGACTG 1272  
| | | | |  
QY 2067 ATCCAGCAGCGCAATCTGCCCCAACAATCCGGCGCAGACGTGGCCAACTCCCAATTCGTCG 2126  
| | | | |  
Db 1273 ATTGACACGGCAGCAAGAGCGCATTTTCGGGCCAAACATCGCGCAACCGCAAGTCATCG 1332  
| | | | |  
QY 2127 GTTTCGACCGCGTTTTCAGATTTGCTGCGCGCGCAGGGCTCGACGAAGATCCCATGCG 2186  
| | | | |  
Db 1333 CGTTCCACGCGCTCTTTCAAAATAGCACTCCACTTTCCCGATTCTCTGGGGCCACACGCG 1392  
| | | | |  
QY 2187 AAGTTCTCGTTTCATCGATCTGGCGGCAATGAGCGGGCGTGGAACACTTCTCGGCCGAT 2246  
| | | | |  
Db 1393 AAGTGTCTCTTTGTGACTTGGCGGCAATGAACGCGGGCGGATACGCAATCCGCGCAT 1452  
| | | | |  
QY 2247 CGCGACAGCGTATGAGGGTCCGAGATTAAACAAATCGCTGTGGCCCTCAAGGAGTGC 2306  
| | | | |  
Db 1453 CGTCAAACCTCGCATCGAGGGGCGGAGATCAATTAATCTCTGCTGGCCCTCAAGGAGTGC 1512  
| | | | |  
QY 2307 ATTCTGCGTTGGGCAAAACAGTGGCCCACTTTCGCCCTTCCTTCCAAATCTCAACCCAG 2366  
| | | | |  
Db 1513 ATTCGAGCCCTCAGCGCGCAGTCGAGTCACCTTCCTTCCTGGCTCCAAAGTTGACCCAA 1572  
| | | | |  
QY 2367 GTGCTCGCGACTCTGTTCAATTGGCG---AGAAGACGAAGAGTGCATGTATGCGCATGATC 2423  
| | | | |  
Db 1573 GTGCTCGCGACTCTCTTTGTGCGCGCAAGAAGAACAGACTGTCATGATTGCCATGATA 1632  
| | | | |  
QY 2424 TCGCGGGAATTAGTCTCTGCGAGCACAGCTCAACAGCTCGCTATGCGGATCGTGTTC 2483  
| | | | |  
Db 1633 TCGCCATCCATGAGCTGCTGGAGAAATACGCTCAACACTCTACGTTACGACAGGGTT 1692  
| | | | |  
QY 2484 AAGGAGCTGGTGGTCAAGGATATCG 2508  
| | | | |  
Db 1693 AAGGAGCTCATGCCAAGGAAGACG 1717  
| | | | |

RESULT 5

ABL05366  
ID ABL05366 standard; cDNA; 4190 BP.

XX ABL05366;

XX AC ABL05366;

XX XX

DT 26-MAR-2002 (first entry)

XX

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 10580.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical; gene; ss.

KW Drosophila melanogaster.

XX WO200171042-A2.

PN 27-SEP-2001.

XX

PF 23-MAR-2001; 2001WO-US09231.

XX

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX

PA (PEKE ) PE CORP NY.

XX

PI Venter JC, Adams M, Li PMD, Myers EW;

XX

DR WPI; 2001-656860/75.

DR P-PSDB; ABB61263.

XX

PT New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell

PT interactions.

XX

PS Claim 1; SEQ ID NO 10580; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-

CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 4190 BP; 1187 A; 1021 C; 1025 G; 957 T; 0 U; 0 Other;

Query Match 18.3%; Score 564.6; DB 4; Length 4190;

Best Local Similarity 68.1%; Pred. No. 4.2e-134;

Matches 834; Conservative 0; Mismatches 379; Indels 12; Gaps 3;

QY 1296 GCATTGAAGAGGTGGAGCGACTGAAGGAGAAATCGGAGAACGCGCCGACAGGCC 1355

Db 1493 GTAGTGAAGAGGTGAATCGCATGAAGGAGCAGAGGAGAGCGAAGGCTGCCAGGCG 1552

QY 1356 GAGATGAAGGAGGAGAGGTGGCGCTGATGAACCCAGGATCCGGGCNATCAAACTGGGAG 1415

Db 1553 GAACAGCTCCAGGAGAGGATGCACTGCGTCGCAATAATCCGGGGAATCCCAACTGGGAG 1612

QY 1416 ACGCGCGCAAAATGATACGCGAAATATCAGAGCAGCTGGAAATTTGTCCGCT-----GCTC 1469

Db 1613 GTGTGCGTGTGCTGCGGCAATACCGCTCCACCTTGATCTTCTCCATTCGATGCCTG 1672

QY 1470 GATGCCAGGCGGTGCGATGACCATCAGATCAGATGTGCGTGCAGAACGCGTCCCAATTAGC 1529

Db 1673 GATCCCAATGGAGGCACTGTTTCAGCAATATACGGTGTGTGTCGGAAGAACGACCCATGAT 1732

QY 1530 CGCAAGGAGGTCAATCGCAGAGGATCGATGTCTATTTCCGTCGCGGCAAGACATGCTC 1589

Db 1733 CGCAAGGAGGAGAACTCCAGAACCTCGACATCATCAGATTTCCAGTTCGCGACAGCCTG 1792

1590 ATCGTGCACGACCGCGCAGCAAGGTTCGACCTCACCAAGTTCTCTGGAGAACCAAGTTT 1649  
1793 ATCGTCCATAGTTGGCTCAAGTGGATCTCACCAAGTTCTCTGGAGAACCAAGTTT 1852  
1650 CGCTTCGACTAGCTTCAAGCAACGTCGCGCAATGCGCATGATATACCAATACACAGCC 1709  
1853 CGTTTCGACTACAGTTCGACGAGAGTCTCAATGCGCTGCTACGATCACACTGCT 1912  
1710 AAGCCGTTGGTGAACCAATTTCCAGGGCGGAATGGGAGCTGCTCGCTACGCCAG 1769  
1913 CGTCCGTTGATCAGAACCATGTTCCAGGGCGCAATGCGCATGTTGTTCTGCTTACGGCAA 1972  
1770 ACGGATCGGCGCAACGACACCATGCGCGGTGAGTTTAATGGAAGGTGCGAGACTGC 1829  
1973 ACTGCGAGCGGAACCAACGACACCATGCGCGGAGATTTCTCGGAAGGTTGAGATTGC 2032  
1830 AAGAACGGGATCTAGCCATGCGCGGCAAGGATGCTTTGTGACCCCTGAATATGCCGCT 1889  
2033 GGTACCGGGATCTAGCCATGCGACTGCGATGCTTTCGAGGAGGTATCGCGCCGAG 2092  
1890 TACCGCGCATGAATCTAGTCTGTCGCGCAGTTTCTTTGAGATTTACAGTGGCAGTTC 1949  
2093 TACCGGCAATGGGTGCGCAAGATTACGTGCGAGTCTTTCGAAATCTATGGCACCAGGTG 2152  
1950 TTGCATC---TTCTGTCCGACAGCAAGAACTGCGCGTCTCGGAGATGGTAAACAGCAA 2006  
2153 TTGCATCTTGTCTACCCCAACAGCCCATGCTCGGGTCTTAGAGATGCGCAGGAG 2212  
2007 GTGCAGGTGGTGGGACTCACCGAAGGTGTCGATGCGCTCGAGGAGGTACTGAAGTTC 2066  
2213 GTCTGTGTGTGGGCTTAACGAGATGCGGTGACCAAGTGGAGATGCTCTGAGACTG 2272  
2067 ATCCAGCAGCGAATGCTGCGCGACATCGCGCAGAGCTGCGCAATCTCAATTCGTCTG 2126  
2273 ATTGAGCAGCGAGCAAGAGCGCACTTCCGGCCAAACATCGCGGAAACGCGCAAGTCTCG 2332  
2127 CGTTTCGACGCGCTTTTCCAGATTGTCTGCGGCGCGCAGGAGTCTGACGAAGATCCATGGC 2186  
2333 CGTTCCACGCGCTTTTCAANTAGCATCTCACCTTCCGATTTCTGGGGCCACACGCG 2392  
2197 AAGTTCCTGATCATGATCTCGCGGCAATGAGCGGGCGTGGACACTTCTCGGCCGAT 2246  
2393 AAGTCTCTTTGTGGACTTGGCGGCAATGAACGCGGGCGGATACGCAATTCGCCGAT 2452  
2247 CGGAGACGCTATGAGGCTGCGGAGATTAAACAATCGCTGCTGGCCCTCAAGAGTGC 2306  
2453 CGTCAAACTCGATCGAGGAGCGGAGATCAATAAATCTCTGCTGGCCCTCAAGGAGTGC 2512  
2307 ATTCTGTGTTGGGCAACAGTCGCGCCCACTTGGCCCTTCCGCTGTCTCCAAACTCACCCAG 2366  
2513 ATTCGAGCCCTCAGCGCGCATCGAGTCACCTTCCCTTCGTTGGCTCGAAGTTCACCCAA 2572  
2367 GTCTGCGCGCATCTGTTCAITTTGGG---AGAAGAGCAAGAGCTGATGATAGCATGATC 2423  
2573 GTGTGCGCGACTCTTTGTGCGGGCAAGAAAGAACAGACCTGATGATTGCGCATGATA 2632  
2424 TCGCGGAGACTTACTCTCTGCGACACACGCTCAACAGCTGCGCTATGCGGATCGGTGTC 2483  
2533 TCGGCATCCATGAGCTGCTGGAGAAATACGCTCAACACTCTACGTTAGCGAGACAGGTT 2692  
2484 AAGGAGCTGTGGTCAAGGATATCG 2508  
2593 AAGGAGCTCATAGCCAGGAGACG 2717

RESULT

ABL05833 standard; cDNA; 1970 BP.

XX ABL05833;

AC ABL05833;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 11981.  
DE Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.  
KW Drosophila melanogaster.  
OS WO200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US009231.  
XX 23-MAR-2000; 2000US-0191637P.  
XX 11-JUL-2000; 2000US-00614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI WPI; 2001-656860/75.  
XX P-PSDB; ABB61730.  
XX New isolated nucleic acid detection reagent for detecting 1000 or more  
genes from Drosophila and for elucidating cell signaling and cell-cell  
interactions.  
XX Claim 1; SEQ ID NO 11981; 21pp + Sequence Listing; English.  
XX The invention relates to an isolated nucleic acid detection reagent  
capable of detecting 1000 or more genes from Drosophila. The invention is  
useful in developmental biology and in elucidating cell signalling and  
cell-cell interactions in higher eukaryotes for the development of  
insecticides, therapeutics and pharmaceutical drugs. The invention  
discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-  
ABBS72072). The sequence data for this patent did not form part of the  
printed specification, but was obtained in electronic format directly  
from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 1970 BP; 515 A; 521 C; 538 G; 396 T; 0 U; 0 Other;  
XX

Query Match 16.4%; Score 505.8; DB 4; Length 1970;  
Best Local Similarity 64.7%; Pred. No. 3.9e-119;  
Matches 822; Conservative 0; Mismatches 432; Indels 17; Gaps 4;  
Qy 1252 CCACAAACCCAGGAGCTGGCGCGCTAGTAGCCCGCGCATCGCACATTTGAAAGAGGTGG 1311  
Db 399 CGGCCACGCCAGCAGCGCGCGCCAGTTTCGAGAGGA--CGAGTGTGTGCCAGGCTG 456  
Qy 1312 AGGACTGAAGGAGATTCGGAGAGCGACGCGCCCGACAGGCGCGATGAGAGAGAGA 1371  
Db 457 AAAGAATGCGAAAAGAAACGGGAAACGGAGGCGAGAGCAAGCAAGCCAGGACTCGTCTAGATC 516  
Qy 1372 AGGTGGCGCTGATGAACACGAGATCCGGCAATCCAAACTCGGAGACGCGCAATGATAC 1431  
Db 517 GGGAGCAGGGGAGAGACGAAGATCCGGGAATCCCAACTGGAGTAGCCAGATGATAC 576  
Qy 1432 GCGAATATCAGAGCACGCTGGAATTTGTGCGCTGCTCGATGG-----CCAGGCCG 1482  
Db 577 GACTGCAACCGAGCAAAATGGAGATCAGCGGTGAGAGTGTGAGAGTGTACTACGAACGA 636  
Qy 1483 TCGATGACCATCAGATCAGATGTGCGTGCAGAGCGTCCATTAGCCCGCAGAGAGTCA 1542  
Db 637 TCAATTGCCACCAAAATTATGTTTGTGTGAGGAGAGACCACTGAGCGCAAGAGGTGG 696  
Qy 1543 ATCGAAGGAGATCGATGTCTATTTCGTTGCGCGCGCAGGACATGCTCATGTGCACGAGC 1602  
Db 697 CTGACCGGGAACAGAGATGTGTCAGCATTCGCTTAGCAACACATTTGTTGTTCCAGAGC 756  
Qy 1603 CGCGCAGCAAGGTGCGACCTTCCCAAGTTCTCGGAGAACCAAGTTTCGTTTCGACTACG 1662



Db 757 CCGCAAGCATGTGAACCTGGTCAAGTTCTCTGAAATCATAGCTTCCTGTTGATTACG 816  
 QY 1663 CTTTCAACAGCAGTGGCGCAATATGCGCATGTATACAAATACACAGCAAGCCGTTGGTGA 1722  
 Db 817 TCTTCGACGAGGAGTCTCCAATGCCACGGTCTACGAAATTCACAGCCGACCTTGATAA 876  
 QY 1723 AAACCATTTTCGAGGCGGAATGGCGACGTGTTCCGCTTACGGCCAGAGCGGATCGGCA 1782  
 Db 877 AGCACATTTTGTATGGCGGAATGGCCACGTGTTTCCCTTACGGACAAACTGGAAGCGGCA 936  
 QY 1783 AAACGCACACCATGGCGGTGAGTTTAAATGGAAGGTGAGGACTGCAAGAACGGCATCT 1842  
 Db 937 AGACCTATACGATGGGTGTCAGTTCCCGGAAGGATCAGAGCTCAATGATGGCATCT 996  
 QY 1843 AGCCATGGCGGCAAGGATGCTTTGTGACCTGCAATATGCGCGTTTACCGCGCATGA 1902  
 Db 997 ATGCAATGGCGCTAAGGACGTGTTCTCCACTTAAAGACGGTTCCTATACAAAGCTTA 1056  
 QY 1903 ATCTAGTCTCTGGCCAGTTCTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTC--- 1959  
 Db 1057 ATCTGAAAGTTTACTGCACTTCTTCGAGATCTACGGCACCCCGGGTGTTCGATCTCTGA 1116  
 QY 1960 TGTCCGACAAAGCAGAACTGCGCTCTCGAGGATGTTAAACAGCAAGTGCAGGTGGTGG 2019  
 Db 1117 TGCCTGGCAAGCCAACTGCGTGTCTTGAGATAGAAACAGCAGGTGCAAGTGGTG 1176  
 QY 2020 GACTCACCGAAGGTGTCGATGGCTGAGAGGTACTGAAGCTCATCCAGCAGCGCA 2079  
 Db 1177 GCCTCACCCAGAATCCAGTACAGAACACCGCGAAGTTCTGACCTACTCGAGTTGGCA 1236  
 QY 2080 ATGCTGCCGACATCCGCGCAGAGTGGGCCAATCCAAATGCTGGGTTCGCGAGCGG 2139  
 Db 1237 ATAGTGTCCGAACCTCGGGTCAACCTCTGCCAATTCGAAGTCTCCCGATCGCATGCTG 1296  
 QY 2140 TTTTCCAGATGTCGTGCGCGCGCAGGGCTCGACGAAGATCCATGCGCAAGTTCCTGTCA 2199  
 Db 1297 TGTTCCAATCTGTCTGAGATCCGCGGGGCGAGACTACCGGAAATTCCTGCTTA 1356  
 QY 2200 TCGATCTGGCGGCAATAGCGGGGCTGGACACTTCTCGCCGATCGGCAAGCGTA 2259  
 Db 1357 TAGATCTGGCGGGAATGAAGAGGAGCGGACACAGCTCGCGGATCGACAGCGGCC 1416  
 QY 2260 TGGAGGTGCGGAGATTAACAATCGTCTGGCCCTCAAGAGTGCAATCGTGGTGG 2319  
 Db 1417 TGGAGGATCCGAGATCAATAAATCGTCTGCTGCTCAAGGAATGCAATCGCGTCTGG 1476  
 QY 2320 GCAACAGTCCGCCCACTTGGCCCTTCCGTGCTCCAACTCACCCAGGTGCTGCGCGACT 2379  
 Db 1477 GCGCGAGTCTGAGTCAATTTGGCAATTCGTTGGCTCCAGTGACCCAAAGTCTGCGGGACT 1536  
 QY 2380 CTTTCATTTGGCG---AGAAGAGCAAGACGTGCAATGATAGCCATGATCTCGCGGACTTA 2436  
 Db 1537 CTTTCATCGGAGTAAGAGGTGAACCTGATGATAGCCATGATCTCGCATGCTTGC 1596  
 QY 2437 GCTCTCGGAGCAGCGCTCAACAGCTGCGGTATGCGGATGCTGTCAAGAGGTGGTGG 2496  
 Db 1597 ATTCTGGTGGAGCATCTTGAACACGCTGCGTTATGCGGATCGGTTGAAGGAACCTAAGT 1656  
 QY 2497 TCAAGGATATC 2507  
 Db 1657 TGGAGTCGATC 1667

RESULT 7  
 ID ABL05832/c  
 XX ABL05832 standard; cDNA; 3970 BP.  
 AC ABL05832;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 11978.  
 XX

KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical; gene; ss.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US009231.  
 XX  
 PR 23-MAR-2000; 2000US-0191637P.  
 PR 11-JUL-2000; 2000US-00614150.  
 XX  
 PA (PEKE) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI; 2001-656860/75.  
 DR P-PSDB; ABB61729.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signaling and cell-cell  
 PT interactions.  
 XX  
 PS Claim 1; SEQ ID NO 11978; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-  
 CC ABB72072). The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 3970 BP; 969 A; 935 C; 968 G; 1098 T; 0 U; 0 Other;  
 Query Match 16.4%; Score 505.8; DB 4; Length 3970;  
 Best Local Similarity 64.7%; Pred. No. 5.6e-119;  
 Matches 822; Conservative 0; Mismatches 432; Indels 17; Gaps 4;  
 QY 1252 CCACAACCCAGGAGCTGGCGGCTAGTACCCGCGCATCGCACCATTTGAAAGAGGTGG 1311  
 Db 2572 CGGCCACGCGCAGCAGCGCGCGCGCATTTGAGAGGA--CGAGTGTGTGACAGGTG 2515  
 QY 1312 AGCGACTGAAGGAGAAATCGCAGAGGACGCGCCCGACAGGCCGAGATGAAGAGGAGA 1371  
 Db 2514 AAAGAATCGGAAAAGAACGGGACGAGGCGGAGAACAGCCAGGACTCGTCTAGATC 2455  
 QY 1372 AGTGGGCGCTGATGACACGAGATCGGCGCATCCAACTGGGAGACGCGGCAATGATAC 1431  
 Db 2454 GGGAGCAGGGGAGAACGAAATCCGGGAAATCCCACTGGGAAAGTAGCCAGAAATGATAC 2395  
 QY 1432 CGGAATATCAGACACGCTGGAATTTGTGCGCTGCTCGATGG-----CCAGGCGG 1482  
 Db 2394 GACTGCAACGCGGACCAATGGAGAGTCAAGGCTGAGAGTGTACTACGACGACGAA 2335  
 QY 1483 TCGATGACCATCAGATCAGTGTGCGTGGCAAGCGTCCATTAGCCGCAAGGAGGTCA 1542  
 Db 2334 TCAATTCGACCAAAATATGTTGTTGTGAGGAGAGACCACTGAGGCGCAAGGAGTGG 2275  
 QY 1543 ATCCGAGGAGATCGATGTCAATTTTCGGTGGCGCGCAGGACATGCTCATCTGCGACGAC 1602  
 Db 2274 CTGACCGGGAAACAGGATGTGGTCAGCATTCCTTAAGCACACATTTGGTGGTCCACGAC 2215  
 QY 1603 CGCGCAGCAAGGTGCACTCACCAAGTTCTTGGAGAACCAAGTTTCGCTTCGACTACG 1662  
 Db 2214 CCGCAAGCATGTGAACCTGCTCAAGTTCCTGGGAAATCATAGTTCGTTTCGATTACG 2155  
 QY 1663 CTTTCAACGACACGTGCGCAAAATGCCATGGTATACAAATACACAGCCAGCCGTTGGTGA 1722

Db 2154 TCTTCGACGAGAGTGCTCCCAATGCCAGGTCACGAATTCACAGCCGACCCCTTGATAA 2095  
Qy 1723 AAACCATTTTCGAGGCGGAATGGCGAGTGTCTTCCTACGCGCAGACGGATCGGCA 1782  
Db 2094 AGCAATTTTGTATGGCGGAATGGCCAGTGTCTTCCTACGCGCAGCAAACTGGAGCGGCA 2035  
Qy 1783 AAACGACACACCATGGCGGCTGAGTTTAATGAAAGGTCAGAGCTGCAAGAACGCGCATCT 1842  
Db 2034 AGACCTATACGATGGTGTGTCAGTTCCCGGAGGACATCAGAGCTCATGATGGCATCT 1975  
Qy 1843 AGCCATCGCGGCGCAAGAGTGTCTTTGTGACCCCTGGAATATGCGCGTTTACCGCGCCATGA 1902  
Db 1974 ATGCAATGGCGCGCTAAGGACGCTGTCTCCACTCTAAAGACGCTTCCCTATAACAAGCTTA 1915  
Qy 1903 ATCTAGTCTGTCTCGGCGAGTTCTTTGAGATTTACAGTGGCAAGGTCTTCATCTTC--- 1959  
Db 1914 ATCTGAAAGTTTACTGTGAGCTTCTTCGAGATCTACGGCACCCTGGGTGTTCATCTCTCTGA 1855  
Qy 1960 TGTCCGACAGCAAGAACTGCGCGCTCTGAGGATGTTAAACAGCAAGTGCAGGTGGTGG 2019  
Db 1854 TGCCTGGCAAGCCACAACCTGCTGTCTTGGAGATAGAAACAGCAGGTGCAAGTGGTGG 1795  
Qy 2020 GACTCACCGAAGGTGTGATGCGTTCGAGGAGGTACTCAAGTCTATCCAGCACGCGCA 2079  
Db 1794 GCCTCACCGAATCCAGTACAGAAACACCGCGGAAGTCTCGACCTACTCGAGTTGGGCA 1735  
Qy 2080 ATGTCGCCGGAACATCCGCGCAGAGCTGCGGCAACTCCAAATTCGCGGTTCGACGCGG 2139  
Db 1734 ATAGTGTCCGAACCTCGGGTCTCACACCTCTGCCAATTTCCAACTCTCCGATCGCATGCTG 1675  
Qy 2140 TTTTCCAGATTTGTCTGCGCGCGCAGGCTCGACGAAGATCCATGGCAAGTTCTCGTTCA 2199  
Db 1674 TGTTCCAATCTGCTGAGATCCGCGCGCGGCGAGAGCTACACGGGAAATTCCTGCTTA 1615  
Qy 2200 TCGATCTGGCGGCAATGAGCGGCGCGTGGACATTTCTCGGCCCATCGGCAGACGCTA 2259  
Db 1614 TAGATCTGGCGGGAATGAAAGAGAGCGGACAAACAGCTCGCGCGATCGACAGACGCGCC 1555  
Qy 2260 TGGAGGTGCGGAGTAAACAAATCGTGTGCGCCCTCAAGAGTGCATTCGTGCGTTGG 2319  
Db 1554 TGGAGGATCCGAGATCAATAAATCGTGTGCTGCTCCCAAGGAATGCAATTCGCGCTCTGG 1495  
Qy 2320 GCAACAGTGGCGCCACTTGGCTCTCGCTGCTCCAAACTCACCAGAGTGTGCGCGACT 2379  
Db 1494 GCGCCAGTTCAGTCAATTTGCAATTCGCTGCTCCAGCTGACCCAAAGTCTCTGCGGACT 1435  
Qy 2380 CGTTCATTTGGCG---AGAGAGCAAGAGTGCATGATAGCCATGATCTGCGCGGACTTA 2436  
Db 1434 CTTTCATCGGAGGTAAGAAAGTGAAGAACTCGATGATAGCCATGATCTCGCCATGCTTGC 1375  
Qy 2437 GCTCTCGGAGCACACGCTCAACAGCTGCGCTATGCGGATCGTGTCAAGGAGCTGGTGG 2496  
Db 1374 ATTGCTGGAGCATACCTTTGAACAGCTGCGTTATGCGGATCGGGTGAAGGAACTAAGTG 1315  
Qy 2497 TCAAGGATATC 2507  
Db 1314 TGGAGTCGATC 1304

RESULT 8  
AAD24080  
ID AAD24080 standard; DNA; 1292 BP.  
AC AAD24080;  
XX  
XX  
XX 09-APR-2002 (first entry)  
DT Human mitotic centromere-associated kinesin protein fragment #1 DNA.  
DE Human; MCAK; mitotic centromere-associated kinesin; motor domain;  
KW cellular proliferation disorder; cancer; hyperplasia; restenosis;  
KW cardiac hypertrophy; immune disorder; inflammation; autoimmune disease;

arthrititis; graft rejection; inflammatory bowel disease; ds.  
XX Homo sapiens.  
XX Key Location/Qualifiers  
FH CDS 1..1292  
FT /\*tag= a  
FT /product= "MCAK enzyme fragment"  
FT /transl\_except= (pos:4..5, aa:Arg)  
XX US6331424-B1.  
PN 18-DEC-2001.  
XX 15-JUN-2000; 2000US-00594669.  
XX 20-APR-1999; 99US-00295612.  
PR 18-MAY-1999; 99US-00314464.  
XX (CYTO-) CYTOKINETICS INC.  
XX Beraud C, Sakowicz R;  
XX WPI; 2002-089075/12.  
DR P-PSDB; AAE14499.  
XX New human MCAK (mitotic centromere-associated kinesin) protein useful in  
PT identifying agents for use in the treatment of cellular proliferation  
PT disorders.  
XX Disclosure; Fig 1; 44pp; English.  
XX The invention relates to human MCAK (mitotic centromere-associated  
CC kinesin) protein, and its fragments that comprise a motor domain and  
CC directly or indirectly produce ADP or phosphate. The MCAK enzyme and its  
CC fragments are used in methods to identify compounds that modulate their  
CC activity. Modulators of MCAK are useful as therapeutic agents for  
CC treating cellular proliferation disorders such as cancer, hyperplasias,  
CC restenosis, cardiac hypertrophy, immune disorders, inflammation,  
CC autoimmune disease, arthritis, graft rejection, inflammatory bowel  
CC disease. The present sequence is a human MCAK fragment encoding DNA  
XX Sequence 1292 BP; 374 A; 287 C; 349 G; 282 T; 0 U; 0 Other;  
SQ  
Query Match 14.8%; Score 454.6; DB 6; Length 1292;  
Best Local Similarity 62.6%; Pred. No. 4.8e-106;  
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;  
Qy 1300 TGAAGAGGTGGAGCGGCTGAAGGAGATCGGAGAGCGACGCGCCGACAGCGCGAGA 1359  
Db 22 TGAAGGAGGTGGAAAAAATGAAGAACAAAGCGAGAGAGAGAGGCGCCAGAACTCTGAAA 81  
Qy 1360 TGAAGGAGGAGAGGTGGCGCTGATGAACGAGATCGGGCAATCCAACTGGGAGACGG 1419  
Db 82 TGAGAAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTCTTTCCAAACCTGGAAATTG 135  
Qy 1420 CGCAAAATGATACGCGAATATCAGAGCAGCGCTGGAATTTGTCGCTGCTCGATGCGCCAGG 1479  
Db 136 CCCGAATGATTAAGAAATTTCCGGGCTACTTTTGGAAATGTCATCCACTTACTATGACTATC 195  
Qy 1480 CCGTCGATGATACAGATCAGATGTCGTCGCGCAAGCGTCCCATTTAGCGCAAGGAGG 1539  
Db 196 CTATCGAGAGCAGAGATATGCTGTGTGTAGGAAACGCCCACTGAAATAGCAAGAT 255  
Qy 1540 TCAATCGCAAGGAGATGATGTCATTTTCGGTTCGCGCAAGGACATGCTCATCTGTCACG 1599  
Db 256 TGGCCAAAGAAAGAAATTTGATGTCATTTCCATTCCTAGCAAGTGTCTCTCTTGTGATATG 315  
Qy 1600 AGCCGCGCAGCAAGGTGCGCTCACCAGTTCCTCGGAGAACCAAGTTTCGCTTCCACT 1659  
Db 316 AACCCAAAGTTGAAAGTGAGACTTAACAAAGTATCTGGAGAAACCAAGCAATTCGCTTTGACT 375  
Qy 1660 ACGCCTTCAACGACACGTCGCGCAATGCGCATGCTGATATCAAAATACACAGCAAGCGCTGG 1719

```
Db 376 TTGCAATTTGATGAACAGCTTCGAATGAAGTTGTCTACAGTTTCACAGCAAGCCACTGG 435
QY 1720 TGAACCACTTTTCGAGGGCGGAATGCGACGTGCTTCGCTACGGCCAGACGGGATCGG 1779
Db 436 TACAGACAATCTTTGAAGGTGGAAGCAACTTTGTTGATATGCGCAGACAGGAAGTG 495
QY 1780 GCAAAAGCAGACACATGCGGGCGGTGAGTTTAATGGAAGGTGCAAGACTGCAAGAACGGCA 1839
Db 496 GCAAGACACATATATGCGGGCGGAGACCTCTCTGGAAGAGCCAGAAATGCATCAAAAGGGA 555
QY 1840 TCTACGCCATGCGGGCGCAAGAGTCTTTGTGACCTGCAATATGCGCGTTACCGCGCCA 1899
Db 556 TCTATGCGATGCGCTCCGGGAGCTCTTCTCTGGAAGATCAACCTGCTACCGGAAGT 615
QY 1900 TGAATCTAGTCTGCTCGGCCAGTTTCTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTC 1959
Db 616 TGGGCGCTGGAAGTCTATGTGACATCTCTCGAGATCTACATGGGAAGCTGTTTGACTGC 675
QY 1960 TGTCCGACAGAGCAAACTGCGCGTCTGAGGATGTTAAACAGCAAGTGCAGGTGGTG 2019
Db 676 TCAACAAGAGGCAAGCTGCGCGTGTGAGGATGCGCAAGCAACAGGTGCAAGTGGTG 735
QY 2020 GACTCACCGAGAGGTGCTGATGGGTGAGGAGTACTGAGCTCATCCAGCGCA 2079
Db 736 GCGTGCAGGAGCATCTGGTTAACTCTGCTGATGTCTCAAGATGATCGACATGGCA 795
QY 2080 ATGCTGCCGGAACATCCGCGCAGACGTGCGCCAACTCCAAATTCGTGCGGTTCCGACCGG 2139
Db 796 GCGCTGCAGAACTCTGGGCGAGACATTTGCCAACTCCAAATTCCTCCGCTCCACGCGT 855
QY 2140 TTTTCAGATTTGCTGCGGCCGCGAGGTGCGAGAGATCCATGCGCAAGTTCTCTTCA 2199
Db 856 GCTTCCAAATTAATCTTCGAGCTAAAG-----GAGAAATGCATGCGCAAGTTCTCTTGG 909
QY 2200 TCGATCTGCGGCAATGAGCGGCGTGGACACTTCTCGGCCGATCGGACACCGTA 2259
Db 910 TAGATCTGGCAGGAATGAGCGGCGCGACACTTCCAGTGTGACCGGCGAGACCGCA 969
QY 2260 TGGAGGTGCGGAGATTAACAAATCGTCTGCGCCCTCAAGGAGTGCATTCGTGCGTGG 2319
Db 970 TGGAGGGCGGAGAAATCAACAGAGTCTCTTAGCCCTGAAGAGTGCATCAGGGCCCTGG 1029
QY 2320 GCAACAGTGGCGCCACTTGCCTTCGCTGTCTCCAAATCTCAACCCAGGTGCTGCGGACT 2379
Db 1030 GACAGAAACAAGGTCAACCCCGTTCGCTGAGAGCAAGTGCACAGGTGCTGAGGACT 1089
QY 2380 CATTCAATGGCGAGAGACAGACGTGATGATAGCCATGATCTCGCCGGGACTTAGCT 2439
Db 1090 CTTTCAATGGGGAGAACTTAGGACTTGCATGATTCGCCAGATCTCACCAGGCATTAAGCT 1149
QY 2440 CTTGCGAGCACACGCTCAACACGCTGCGCTATGCGGATCGTGTCAAGGAGCTG 2492
Db 1150 CTTGTGAATATACTTTAAACACCTGAGATATGACAGAGGTCAAGGAGCTG 1202
```

## RESULT 9

ADE10082

ID ADE10082 standard; DNA; 1292 BP.

XX AC ADE10082;

XX AC ADE10082;

XX DT 29-JAN-2004 (first entry)

XX DE DNA encoding human MCAK enzyme fragment #1.

XX KW Human; motor protein; mitotic centromere-associated kinesin; MCAK;  
KW microtubule-stimulated ATPase activity; cellular proliferation disorder;  
KW cancer; autoimmune disease; arthritis; inflammatory bowel disease;  
KW keloid; psoriasis; tumour; cytostatic; immunosuppressive; antiarthritic;  
KW antiinflammatory; gastrointestinal; vulnery; antipsoriatic; ds.

XX OS Homo sapiens.

XX FH

FT CDS

FT

FT

FT

FT

FT

XX

PN

XX

XX

PD

XX

PF

XX

PR

XX

PA

XX

PI

XX

DR

DR

XX

PT

PT

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

Location/Qualifiers

1..1292

/tag= a

/transl\_except= (pos:4..5,aa:Arg)

/note= "This codon has an apparent 1 nucleotide deletion which alters the reading frame"

US6638754-B1.

28-OCT-2003.

29-MAR-2002; 2002US-00112432.

28-NOV-2000; 2000US-00724215.

(CYTO-) CYTOKINETICS INC.

Beraud C, Sakowicz R;

WPI; 2003-842789/78.

P-PSDB; ADE10083.

New nucleic acid encoding motor proteins, useful in identifying compounds for the treatment of cellular proliferation disorders, such as cancer, autoimmune disease, inflammatory bowel disease, arthritis, keloids and psoriasis.

Claim 2; Fig 1; 44pp; English.

The present invention relates to the isolation of polynucleotide sequences encoding fragments of human motor protein, mitotic centromere-associated kinesin (MCAK). The MCAK enzyme fragments have microtubule-stimulated ATPase activity. The invention also discloses methods for the use of the motor protein fragments. The methods and compositions are useful for high throughput screening systems for identifying compounds useful in the treatment of cellular proliferation disorders, such as cancer, autoimmune disease, arthritis, inflammatory bowel disease, keloids, psoriasis and tumours. The present sequence encodes a human MCAK enzyme fragment.

Sequence 1292 BP; 374 A; 287 C; 349 G; 282 T; 0 U; 0 Other;

Query Match

Best Local Similarity 14.8%; Score 454.6; DB 10; Length 1292;

Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;

QY 1300 TGAAGAGGTGAGGCGACTGAAGGAGATCGCGAGAGCGACGCGCCGACAGGCGGAGA 1359

Db 22 TGAAGGAAGTGGAAAAAATGAAGAACAGCGAGAGAGAGGAAGGCCAGAACTCTGAAA 81

QY 1360 TGAAGGAGGAGAGGTGGCGCTGATGAACAGGATCCGGGCAATCCAAACTGGGAGACGG 1419

Db 82 TGAAGATGAAGAGAG-----CTCAGGAGTATGACAGTAGTTTTCNAACTGGGAATTG 135

QY 1420 CGCAATATGATACGGAATATCAGAGCAGCTGGAATTTGTGCGGTGCTCGATGGCCAGG 1479

Db 136 CCGGAATGATTAAAGAAATTTTCGGGCTACTTTGGAATGTCTACCTTACTATGACTGATC 195

QY 1480 CCGTGTGATGACCATCAGATCAGGTGCGTGGCGGAGGTCCTCATTAGCCGCAAGAGG 1539

Db 196 CTATCGAAGAGCAGAAATATGTCTGTGTGTTAGGAACGCCCACTGAATAAGCAAGAAT 255

QY 1540 TCAATCGCAAGGAGATCGATGTCATTTGCGTCCGCGAAGGACATGCTCATCTGTCACG 1599

Db 256 TGGCCAGGAAGAAATGATGATTTTCAATTCCTAGCAAGTGTCTCTCTTGGTACATG 315

QY 1600 AGCCGCGCAGCAGGTGCGACCTCACCAGTTCTCTGGAGAACCAAGTTTTCGCTTCGACT 1659

Db 316 RAACCAAGTTGAAGTGGGACTTTAAACAAGTATCTGGAGAACCAAGCATTTCTGCTT 375

QY 1660 AGCCCTTCAACGACACGTCGCGCAATGCCATGCGTATACAAATACACAGCAGCCGTTGG 1719

Db 376 TTGCATTTGATGAACACAGCTTCGAATGAAGTTGTCTACAGGTTTACAGCAAGCCACTGG 435  
Qy 1720 TGAACCAATTTTCGAGGCGCAATGGCGACGTGCTTCGCCCTACGCGCAGACGGGATCGG 1779  
Db 436 TACAGACATCTTTGAGGTGGAAAGCACTTGTGTCATATGCGCAGACAGGAAGTG 495  
Qy 1780 GCAAAACGACACCACTGGCGGTGAGTTTAAATGAAAGGTGCAGGACTGCAAGAACGGCA 1839  
Db 496 GCAAGACACATACTATGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCATCCAAAGGA 555  
Qy 1840 TCTAGCCATGGCGGCAAGGATGCTTTGTGACCCCTGAATGATGCGCGTTACCGCGCA 1899  
Db 556 TCTATGCCATGGCCCTCCGGGAGCTCTTCTCTGAAGAATCAACCTGCTACCGGAAGT 615  
Qy 1900 TGAATCTAGTCTGCTCGCCAGTTCTTTGAGATTTTACAGTGGCAAGCTCTGCATCTTC 1959  
Db 616 TGGGCTGGAAGTCTATGTGACATCTTCGAGATCTACAATGGGAAGCTGTTGACCTGC 675  
Qy 1960 TGTCCGACAAGCAGAAATCGCGCTCTGGAGGATGTTAAACAGCAAGTGCAGGTGGTGG 2019  
Db 676 TCAACAAGAAGCCAAAGCTGCGCGTCTGGAGATGGCAACACAGGTGCAAGTGGTGG 735  
Qy 2020 GACTCACCAGAGAGTGTGATGCGCTGCGAGGAGTACTGAAGTCTATCAGACAGGCA 2079  
Db 736 GGTCTGAGGAGCATCTGTTAACTCTGCTGATGATGTCAATCAAGATGATGACATGGGCA 795  
Qy 2080 ATGCTGCCGCAACATCGCGCAGACGCTCGGCCCACTTCCAAATTCGTCGGTTCGCAAGCG 2139  
Db 796 GCGCTGAGAACCTCTGGGAGACATTTGCCAATCTCTTCCGCTCCACAGCGT 855  
Qy 2140 TTTTCCAGATTTGCTGCGCGCGCAGGCTCGACGAAGATCCATGGCAAGTTCTCGTTCA 2199  
Db 856 GCTTCCAAATATTCTTCGAGCTAAGG-----GAGAATGCATGGCAAGTTCTCTTGG 909  
Qy 2200 TCGATCTGGCGGCAATGAGCGGCGTGGACACTTCTTCGCGCATCGGCGAGCGGTA 2259  
Db 910 TAGATCTGGCAGGGAATGAGCGGCGGACACTTTCAGTGTCTGACCGGCGAGACCCGCA 969  
Qy 2260 TGGAGGTGCGAGATTAACAAATCGCTGCTGGCCCTCAAGGATGCAATTCGTGGTTGG 2319  
Db 970 TGGAGGCGCAGAAATCAACAGAGTCTCTTAGCCCTGAAGGATGCAATCAGGCGCCCTGG 1029  
Qy 2320 GCAACAGCTCGGCCCACTTGCCTTCCGTGTCTCCAAACTCACCCAGGTCTGGCGCACT 2379  
Db 1030 GACGAACAAGGCTCACACCCCGTTCGTGAGAGCAAGCTGACAGAGTGTGAGGACT 1089  
Qy 2380 GTTCATTTGGCGAGAGAGAGAGTGCATGATAGCCATGATCTGCGCGGACTTAGCT 2439  
Db 1090 CTTTCATTTGGGAGAACTCTAGGACTTGCATGATGATGATGATGATGATGATGATGATGAT 1149  
Qy 2440 CCTCGAGCACACGCTCAACACGCTGCGCTATGCGATCGGTGTCAGGAGCTG 2492  
Db 1150 CTTGTAATATCTTTAAACACCCCTGAGATGATGAGAGGCTCAAGGAGCTG 1202

RESULT 10  
AD052609 standard; DNA; 1292 BP.  
XX AD052609;  
AC AD052609;  
XX  
DT 12 AUG-2004 (first entry)  
XX Human mitotic centromere-associated kinesin (MCAK) DNA #1.  
XX  
KW Mitotic centromere-associated kinesin; MCAK;  
KW cellular proliferative disease; cancer; hyperplasia; restenosis;  
KW cardiac hypertrophy; immune disorder; inflammation; cytostatic;  
KW vasotopic; therapy; human; gene; ds.  
OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 1. .1292  
FT /\*tag= a  
FT /product= "Human MCAK protein #1"  
FT /transl\_except= (pos:4. .5,aa:Arg)  
XX  
PN US2004096949-A1.  
XX  
XX 20-MAY-2004.  
XX  
XX 28-AUG-2003; 2003US-00651510.  
XX  
XX 28-NOV-2000; 2000US-00724215.  
XX  
XX 29-MAR-2002; 2002US-00112432.  
XX  
XX (CYTO-) CYTOKINETICS INC.  
XX  
XX Beraud C, Sakowicz R;  
XX  
XX WPI; 2004-389161/36.  
XX  
XX P-PSDB; ADO52610.  
XX  
XX Novel isolated microtubule motor protein useful for identifying candidate agent for treating cellular proliferative diseases such as cancer, hyperplasia or restenosis.  
XX  
XX Claim 3; SEQ ID NO 1; 46pp; English.  
XX  
XX The present invention provides novel mitotic centromere-associated kinesin (MCAK) polypeptides and their encoding polynucleotides. The invention is useful for treating cellular proliferative diseases such as cancer, hyperplasia, restenosis, cardiac hypertrophy, immune disorders and inflammation. The invention also acts as a cytostatic and vasotropic agent. The present sequence is human mitotic centromere-associated kinesin (MCAK) DNA.  
XX  
XX Sequence 1292 BP; 374 A; 287 C; 349 G; 282 T; 0 U; 0 Other;

Query Match 14.8%; Score 454.6; DB 12; Length 1292;  
Best Local Similarity 62.6%; Pred. No. 4.8e-106;  
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;  
Qy 1300 TGAAGAGGTGGAGCGACTGAAGGAGAAATCGGAGAGCGACGCGCCGACAGCCCGAGA 1359  
Db 22 TGAAGAGGTGGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAA 81  
Qy 1360 TGAAGAGGAGAGAGGTGGCGCTGATGAACCAAGATCCGGGCAATCCAAACTGGGAGACGG 1419  
Db 82 TGAGATGAAGAGAGAG-----CTCAGGAGTATGACAGTAGTCTTCCAAACTGGGAATTTG 135  
Qy 1420 CGCAATGATACGCGAATATCAGAGCAGCGCTGGAAATTTGCGCGTCTCGATGCCGAGG 1479  
Db 136 CCGAATGATTAAGAAATTTCCGGGCTACTTTGGAATGTCTATCCACTTACTATGCTGATC 195  
Qy 1480 CCGTCTGATGACCATCAGATCAGATGTCGTGGCGCAAGCGTCCCATTTAGCGCAAGGAGG 1539  
Db 196 CTATCGAAGACACAGAAATATGTCTGTCTTAGGAAACGCCCACTGAATAAGCAAGAT 255  
Qy 1540 TCAATCGCAAGGAGATCGATGTCAATTTTCGGTGGCGGCAAGGACATGCTCATCTGTGACG 1599  
Db 256 TGGCCAAAGAAAGAAATTTGATGTGATTTCCATTTCTAGCAAGTGTCTCTCTTGTGATG 315  
Qy 1600 AGCGCGCAGCAAGGTGCGACTCCACCAAGTTCTTGGAGAACCAACAGTTTCGCTTCGACT 1659  
Db 316 AACCAAGTTGAAGTGGACTTAACAAAGTATCTGGAGAACCAAGCAATTCGCTTTGACT 375  
Qy 1660 ACGCCTTCAACGACACGTGGCAATGCCATGCTGATGATGATGATGATGATGATGATGATG 1719  
Db 376 TTGCAATTTGATGAACAGCTTCGAATGAAGTTGCTTACAGGTTTACAGCAAGGCCACTGG 435  
Qy 1720 TGAACCAATTTTCGAGGCGGGAATGGCGAGCTGCTTCGCTAGCGGCGACAGCGGATCGG 1779  
Db 436 TACAGCAATCTTTTGAAGGTGGAAAGCACTTTGTTTGTGATATGGCGACAGCAAGAGTG 495



```
QY 1780 GCACAAACGACACCATGGCGGTGAGTTTAAATGGAAAGTGTCAGGACTGCAAGACGGCA 1839
Db 496 GCAAGACACATATATGGCGGAGACCTCTCTGGGAAAGCCAGAAATCATCCAAAGGGA 555
QY 1840 TCTACGCATGGCGGCCAAGAGATGCTTTTGTGACCCCTGAATATGCGCGTTACCGCGCCA 1899
Db 556 TCTATGCCATGGCTCCCGGAGCGTCTTCTCCTGAAGAATCAACCTCTGCTACCGGAAGT 615
QY 1900 TGAATCTAGTCTGCTCGGCCAGTTCTTTTGTGATTTACAGTGGCAAGCTCTTCGATCTTC 1959
Db 616 TGGGCTCGAAGTCTATGTGACATCTTTCGAGATCTACAATGGGAAGCTGTTGACCTGCG 675
QY 1960 TGTCCGACAAAGCAGAAATCGCGCTCTGGAGGATGTTAAACAGCAAGTGCAGGTGGTGG 2019
Db 676 TCAACAAGAGGCCAAGCTGCGCGTCTGGAGGATGGCAAGCAAGGTGCAAGTGGTGG 735
QY 2020 GACTCACCAGAGAGTGTGATGCGGTCGAGGAGGTACTGAAGTCTCATCAGCACGGCA 2079
Db 736 GCGTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATGACATGGGCA 795
QY 2080 ATGCTGCCCGACATCCGCGCCAGAGTGGCCCACTCCAATTCGTCGGTTCGACCGCG 2139
Db 796 GCGCTGCAGAACCTCTGGGAGACATTTGCCAACTCCAAATTCCTCCCGCTCCACGCGT 855
QY 2140 TTTTCCAGATGCTGTCGCGCGCGCAGGCTCGACGAAGATCCATGGCAAGTTCTCGTTCA 2199
Db 856 GCTTCCAAATTAATCTTCAGCTAAGG-----GAGATGATGGCAAGTTCTTTGG 909
QY 2200 TCGATCTCGCGGCAATGAGCGGGCGTGGACACTTCTCGGCCGATCGGCGAGACGGTA 2259
Db 910 TAGATCTGGCAGGGAATGAGCGCGGACACTTCCAGTGTGACCGGAGACCCGCA 969
QY 2260 TGGAGGTGCCGAGATTAACAATCGCTGCTGGCCCTCAAGAGTGCATTCGTGCGTTGG 2319
Db 970 TGGAGGGCGCAGAAATCAACAAGAGTCTCTTAGCCCTGAAGGAGTGCATCAGGCGCCCTGG 1029
QY 2320 GCAAAACAGTGGCGCCACTTGCCTTCCGTGCTCCAAACTCACCAGTGTGCGCGACT 2379
Db 1030 GACAGAACAGGCTCACACCCGTTCCGTGAGAGCAAGCTGACAGAGTGTGAGGAGCT 1089
QY 2380 CGTTCAATGGCGAAGAGCAAGAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 2439
Db 1090 CTTTCAATGGGAGAACTCTAGGACTTGCATGATGATGATGATGATGATGATGATGATGATGAT 1149
QY 2440 CTTGGAGCACACCTCAACAGCTGCGCTATGCGGATCGTGTCAAGAGCTG 2492
Db 1150 CCTGTGAATATCTTTAAACACCCCTGAGATATGAGACAGGCTCAAGAGCTG 1202
```

```
RESULT 12
ADE1008
ID AD10988 standard; DNA; 1421 BP.
XX AC AD10988;
XX AC
XX AC
DT 29 JAN-2004 (first entry)
XX DE
DE DNA encoding human MCAK enzyme fragment #4.
XX
XX Human; motor protein; mitotic centromere-associated kinesin; MCAK;
XX microtubule-stimulated ATPase activity; cellular proliferation disorder;
XX cancer; autoimmune disease; arthritis; inflammatory bowel disease;
XX keloid; psoriasis; tumour; cystostatic; immunosuppressive; antiarthritis;
XX antiinflammatory; gastrointestinal; vulnery; antipsoriatic; ds.
XX
OS Homo sapiens.
XX
XX Location/Qualifiers
XX 1..1421
XX /*tag=
XX /trans_except= (pos:4..5,aa:Arg)
XX /note= "This codon has an apparent 1 nucleotide deletion
```

which alters the reading frame"

US6638754-B1.  
28-OCT-2003.  
29-MAR-2002; 2002US-00112432.  
28-NOV-2000; 2000US-00724215.  
(CYTO-) CYTOKINETICS INC.  
Beraud C, Sakowicz R;  
WPI; 2003-842789/78.  
P-PSDB; ADE10089.

New nucleic acid encoding motor proteins, useful in identifying compounds for the treatment of cellular proliferation disorders, such as cancer, autoimmune disease, inflammatory bowel disease, arthritis, keloids and psoriasis.

Claim 2; Fig 7; 44pp; English.

The present invention relates to the isolation of polynucleotide sequences encoding fragments of human motor protein, mitotic centromere-associated kinesin (MCAK). The MCAK enzyme fragments have microtubule-stimulated ATPase activity. The invention also discloses methods for the use of the motor protein fragments. The methods and compositions are useful for high throughput screening systems for identifying compounds useful in the treatment of cellular proliferation disorders, such as cancer, autoimmune disease, arthritis, inflammatory bowel disease, keloids, psoriasis and tumours. The present sequence encodes a human MCAK enzyme fragment.

Sequence 1421 BP; 417 A; 313 C; 387 G; 304 T; 0 U; 0 Other;

Query Match 14.88; Score 454.6; DB 10; Length 1421;  
Best Local Similarity 62.6%; Pred. No. 5e-106;  
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;

QY 1300 TGAAGAGGTGGAGCGCACTGAAGGAGAAATCCGAGAGAGCGAGCGCCGACAGCCGAGA 1359  
Db 22 TGAAGAGGTGGAAAAATGAAGACAAAGCAGAGAGAGAGAGGCCCGACGACTCTGAAA 81  
QY 1360 TGAAGAGGAGAGAGGTGGCGCTGATGAACCCAGGATCCGGGCAATCCAAATCTGGAGACGG 1419  
Db 82 TGAGATGAAGAGAG-----CTCAGGAGTATGACAGTAGTCTTTCCAAACTGGGAATTTG 135  
QY 1420 CGCAATGATACGGAATATCAGAGCAGCGCTGGAAATTTGTCGCTGCTCGATGCCGAGG 1479  
Db 136 CCGAATGATTAAGAAATTTGGGCTACTTTGGAAATGTCATCACTTACTATGATGATC 195  
QY 1480 CCGTCCGATGACCATCAGATCAGATGTCGTCGTCGCAAGCGTCCCATTTAGCGCAAGAGG 1539  
Db 196 CTATCGAAGAGCAGAGATATGTCGTGTTAGGAAACGCCCACTGTAATGAAGCAAGAT 255  
QY 1540 TCAATCGCAAGGAGATCGATGTCATTTTCGGTGGCGCAAGGACATGCTCATGTCGACG 1599  
Db 256 TGGCAAGAAAGAAATTTGATGTGATTTCCATTCCTAGCAAGTGTCTCTCTTGGTACATG 315  
QY 1600 AGCGCGCAGCAAGGTGCGCTCACCAGATTTCTGGGAGACCAACAGTTTCGCTTCGACT 1659  
Db 316 AACCCAAAGTTGAAGTGAGCTTAACAAAGTATCTGGGAGAACCAAGCAATCTGCTTTGACT 375  
QY 1660 ACGCCTTCAACGACAGCTGCGCAATGCCATGGTATACAATATACAGCAAGCCGTTGG 1719  
Db 376 TTGCAITTTGATGAACAGCTTCGAAATGAAGTTGCTACAGTTTCACAGCAAGGCCACTGG 435  
QY 1720 TGAACACCATTTTCGAGGCGGGAATGGCGAGCTGCTTCGCTACGGCGACAGCGGATCGG 1779  
Db 436 TACAGACAATCTTTGAAGGTGGAAAAAGCACTTTGTTTTCATATATGCGCAGACAGGAAGTG 495

```
QY 1780 GCAAAACGCACACCATGGCGGTGAGTTTAAATGGAAGGTGCAGGACTGCAAGAACGGCA 1839
DB 496 GCAAGACACATATATGGCGGAGAGCCTCTCTGGAAAGCCAGAAATGCAATGCATCANAGGGA 555
QY 1840 TCTACGCCATGGCGGCAAGGATGTCTTTGTGACCCCTGAAATATGCCGCGTTACCGCGCCA 1899
DB 556 TCTATGCCATGGCTCCCGGAGCTCTTCTCTGGAAGATCAACCTGCTACCGGAAGT 615
QY 1900 TGAATCTAGTCTGCTCGGCCAGTTCTTTGAGATTTTACAGTGGCAAGGCTTTCGATCTTC 1959
DB 616 TGGGCGCTGGAAGTCTATGTGACATTTCTTCAGATCTACAATGGGAAGCTGTTTGACCTGC 675
QY 1960 TGTCCGACAGCAGAACTGCGGCTCTGAGGATGTTAAACAGCAAGTGCAGGTGGTGG 2019
DB 676 TCACCAAGAGGCCAAGCTGCGGCTGCTGAGGATGGCAAGCAACAGGTTGCAAGTGGTGG 735
QY 2020 GACTCACCGAAGAGTGGTGCATGGGCTCGAGAGGTACTGAAAGCTCATCCAGCAGCGCA 2079
DB 736 GCGTCGAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATCGACATGGCA 795
QY 2080 ATGCTCCCGCAACATCCGCGCAGACGTGCGGCCAACTCCAAATTCGTCGCTTGCACGCCG 2139
DB 796 GCGCCTGCAAGAACCTCTGGCGAGACATTTGCCAACTCCAATTCCTCCGCTCCCAAGCT 855
QY 2140 TTTTCCAGATTGCTGCGCGCGCAGGCTCGACGAGATCCATGCAAGTTCTGTTCA 2199
DB 856 GCTTCCAAATTAATCTTCGAGCTAAAGG-----GAGAATGCAATGCAAGTTCTCTTTGG 909
QY 2200 TCGATCTGGCGGCAATGAGCGGCGTGGACACTTCTCGCGCGATCGGCAGACGCTA 2259
DB 910 TAGATCTGCGAGGATGAGCGGCGGACACTTCCAGTCTGACCGGCGAGACCGCA 969
QY 2260 TGGAGGTCGCGAGATTAAACAAATCGCTGCTGCGCTCAAGAGTGCATTCGTGCGTTGG 2319
DB 970 TGGAGGCGCAGAAATCAACAGAGTCTCTTAGCCCTGAAGAGTGCATCAGGCGCTGG 1029
QY 2320 GCAACAGTTCGCGCCACTTGGCTCCGTTGCTGCTCAGAACTCACCGAGTGTGCGGACT 2379
DB 1030 GACAGAACAGGCTCACCCCGTTCGCTGAGAGCAAGTGCACAGAGTGTGAGGACT 1089
QY 2380 CGTTCATTGGCGAGAGCAAGACGTGCATGATAGCCATGATCTCGCGGCACTTAGCT 2439
DB 1090 CTTTCATTGGGAGAACTCTAGGACTTGCATGATGTCACAGATCTCACAGGCATAAGCT 1149
QY 2440 CTGCGAGCACAGCTCAACAGCTGCGCTATCGGATCGTGTCAAGGAGCTG 2492
DB 1150 CCTGTGAATATCTTTAAACACCTGAGATATGCAGACAGGCTCAAGGAGCTG 1202
```

## RESULT 13

AD052615  
ID AD052615 standard; DNA; 1421 BP.

AC AC

XX AD052615;

XX AD052615;

DT 12-AUG-2004 (first entry)

DE Human mitotic centromere-associated kinesin (MCAK) DNA #4.

KW Mitotic centromere-associated kinesin; MCAK;  
KW cellular proliferative disease; cancer; hyperplasia; restenosis;  
KW cardiac hypertrophy; immune disorder; inflammation; cytostatic;  
KW vasotropic; therapy; human; gene; ds.

OS Homo sapiens.

PH Key Location/Qualifiers

FT 1. .1421

FT /\*tag= a

FT /product= "Human MCAK protein #4"

FT /transl\_except= (pos:4. .5,aa:Arg)

XX US2004096949-A1.

PN

XX

PD 20-MAY-2004.

XX 28-AUG-2003; 2003US-00651510.

PF 28-NOV-2000; 2000US-00724215.

PR 29-MAR-2002; 2002US-00112432.

XX (CYTO-) CYTOKINETICS INC.

XX Beraud C, Sakowicz R;

XX WPI; 2004-389161/36.

DR P-PSDB; AD052616.

XX Novel isolated microtubule motor protein useful for identifying candidate

PT agent for treating cellular proliferative diseases such as cancer,

PT hyperplasia or restenosis.

XX Claim 3; SEQ ID NO 7; 46pp; English.

XX The present invention provides novel mitotic centromere-associated

CC kinesin (MCAK) polypeptides and their encoding polynucleotides. The

CC invention is useful for treating cellular proliferative diseases such as

CC cancer, hyperplasia, restenosis, cardiac hypertrophy, immune disorders

CC and inflammation. The invention also acts as a cytostatic and vasotropic

CC agent. The present sequence is human mitotic centromere-associated

CC kinesin (MCAK) DNA.

XX Sequence 1421 BP; 417 A; 313 C; 387 G; 304 T; 0 U; 0 Other;

SQ

Query Match 14.8%; Score 454.6; DB 12; Length 1421;

Best Local Similarity 62.6%; Pred. No. 5e-106;

Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;

QY 1300 TGAAGAGGTGGAGGACCTGAAGAGAAATCGCGAGAGGACGCGCCGACAGCCGAGA 1359

DB 22 TGAAGGAAGTGGAAAAAATGAAGAACAAAGCGAGAGAGAGAAAGGCCCAAGAACTCTGAAA 81

QY 1360 TGAAGGAGGAGAGAGTGGCGCTGATGAACAGAGATCCGGGCAATCCAACTCGGAGACG 1419

DB 82 TGAAGATGAAGAGAG-----CTCAGGAGTATGACAGTAGTTTTTCAAACTGGGAATTTG 135

QY 1420 CGCAATGATACGGAATATCAGAGCAGCTGGAATTTGTGCGCTGCTCGATGGCCAGG 1479

DB 136 CCGCAATGATTAAGAAATTTGGGGCTACTTTGGAATGTCTCACTTACTATGACTGATC 195

QY 1480 CCGTGCATGACCATCAGATCAGTGTGGTGGCGCAAGGTCCTCCATTAGCCGCAAGGAGG 1539

DB 196 CTATCGAAGAGCAGAGAAATATGTGTCTGTGTAGGAAACGCCCACTGAATAAGCAAGAAT 255

QY 1540 TCAATCGCAAGGAGATCGATGTCATTTTCGGTGGCGGCAAGGACATGCTCATCGTCAGC 1599

DB 256 TGGCCAAAGAGAAATGATGATGATTTTCATTCCTAGCAAGTGTCTCTCTTGGTACATG 315

QY 1600 AGCCGCGCAGCAAGGTGCACTTCAACAAAGTTCTCTGGAGAACCAAGTTTCGCTTCGACT 1659

DB 316 AACCCAAAGTTGAAAGTGGACTTTAAACAAAGTATCTGGAGAACCAAGCATTTCTGCTT 375

QY 1660 AGCCCTTCAACGACAGCTGCGCAATGCCATGGTATACAAATACACAGCAAGCCGTTGG 1719

DB 376 TTGCATTTGATGAAACAGCTTCGAATGAAGTTGTCTACAGGTTTCACAGCAAGGCCACTGG 435

QY 1720 TGAATAACCATTTTCGAGGCGGAAATGGCGAGTGTCTCGCTTACGCCAGAGCGGATCGG 1779

DB 436 TACAGACAAATCTTTGAAGTGAAGCAACTTTGTCATATGCGCAGACAGGAAGTG 495

QY 1780 GCAAAACGCACACCATATGGCGGTGAGTTTAAATGGAAGGTGAGGACTGCAAGAACGGCA 1839

DB 496 GCAAGACACATACTATATGGCGGAGACCTCTCTGGGAAGGCCAGAAATGCATCACAAGGGA 555

QY 1840 TCTACGCCATCGCGGCCAAGGATGTTTGTGACCCCTGAATATGCGCGGTATTACCGGCCA 1899

DB







QY	2020	GACTCACCAGAGGTGGTCGATGGCGTCGAGAGGTACTGAAGCTCATCCAGCACGGCA	2079
Db	1292	GGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCAATCAAGATGATCGACATGGGCA	1351
QY	2080	ATGCTGCCGGAACATCCGGCCAGAGCTCGGCCAACTCCAAATTCGTGGGTTTCGCACGCCG	2139
Db	1352	GGCCCTGCAGAACCTCTGGGCAGACATTTGCCAACTTCCAAATTCCTCCCGCTCCACAGCGT	1411
QY	2140	TTTTCCAGATTGCTGCGGCCGCGAGGCTCGACGAAGATCCATGGCAAGTTCCTGTTCA	2199
Db	1412	GCTTCCAAATTAATCTTCGAGCTAAAGG-----GAGAAATGCATGCCAAGTTCCTTTGG	1465
QY	2200	TCGATCTGGCGGGCAATGAGCGGGCGTGGACACTTCCTCGGCCGATCGGCAGACGGTA	2259
Db	1466	TGATCTGGCAGGGAATGAGCGAGCGCGGACACTTCAGTGCTGACCGGCAGACCCGCA	1525
QY	2260	TGGAGGGTGCAGAGTTAACTGCTGCGCCCTCAAGGAGTGCATTCGTGCGTTGG	2319
Db	1526	TGGAGGGCGCAGAAATCAACAGAGTCTCTTAGCCCTGAAGGAGTGCATCAGGGCCCTGG	1585
QY	2320	GCAACAGTCCGCCCACTTGCCCTTCGGTCTCCTCAAACTCACCAGGTGCTGCGGACT	2379
Db	1586	GACAGACAAGGCTCACACCCCGTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGGACT	1645
QY	2380	CGTTCAATTGGCGAGAGAGCAAGCGTGCATGATAGCCATGATCTCGCCGGGACTTAGCT	2439
Db	1646	CCTTCATTGGGGAGAACTCTAGGACTTGCAATGATTTGCCAGATCTCACAGGCATAGCT	1705
QY	2440	CCTGCGAGCACGCTCAACAGCTGCGCTATGCGGATCGTGTCAAGGAGCTG	2492
Db	1706	CCTGTGAATATACTTTAAACACCCCTGAGATATGCAGACAGGGTCAAGGAGCTG	1758

Search completed: May 12, 2006, 21:31:17  
 Job time : 1779 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2006, 21:18:27 ; Search time 2381 Seconds  
(without alignments)  
10700.527 Million cell updates/sec

Title: US-10-840-060-142  
Perfect score: 3081  
Sequence: 1 aaactaaaaattgtgtgc.....gccgaatgcaagcgtagt 3081

Scoring table: IDENTITY NUC  
Gapen 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA.Main:\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3081	100.0	3081	9	US-10-840-060-142
2	3081	100.0	3081	10	US-11-097-143-1103
3	1547.8	50.2	8043	10	US-11-097-143-1102
4	564.6	18.3	2190	10	US-11-097-143-5291
5	564.6	18.3	4190	10	US-11-097-143-5290
6	505.8	16.4	1970	10	US-11-097-143-5990
7	505.8	16.4	3970	10	US-11-097-143-5989
8	454.6	14.8	1292	7	US-10-651-510-1
9	454.6	14.8	1421	7	US-10-651-510-7
10	454.6	14.8	2172	7	US-10-651-510-13
11	453	14.7	2865	6	US-10-240-965-208
12	449.8	14.6	1175	7	US-10-651-510-3
13	449.8	14.6	1304	7	US-10-651-510-9
14	449.8	14.6	2740	3	US-09-954-456-77
15	449.8	14.6	2740	3	US-09-954-456-710
16	449.8	14.6	2740	3	US-09-954-456-1158
17	449.8	14.6	2740	3	US-09-960-253-174
18	449.8	14.6	2740	3	US-09-960-253-179
19	449.8	14.6	2740	3	US-09-849-602-5
20	449.8	14.6	2740	7	US-10-651-510-15
21	449.8	14.6	2740	9	US-10-843-641A-3104
22	449.8	14.6	2740	9	US-10-843-641A-3737
23	449.8	14.6	2740	9	US-10-843-641A-4185

ALIGNMENTS

RESULT 1

US-10-840-060-142  
; Sequence 142, Application US/10840060  
; Publication No. US20050227243A1  
; GENERAL INFORMATION:  
; APPLICANT: Cyclacel Limited  
; APPLICANT: Deak, Peter  
; APPLICANT: Frenz, Lisa  
; APPLICANT: Glover, David  
; APPLICANT: Midgley, Carol  
; TITLE OF INVENTION: Cell Cycle Progression Proteins  
; FILE REFERENCE: 10069/2012  
; CURRENT APPLICATION NUMBER: US/10/840,060  
; CURRENT FILING DATE: 2004-05-05  
; PRIOR APPLICATION NUMBER: PCT/GB02/04780  
; PRIOR FILING DATE: 2002-10-23  
; PRIOR APPLICATION NUMBER: GB 0126506.5  
; PRIOR FILING DATE: 2001-11-05  
; PRIOR APPLICATION NUMBER: GB 0128384.5  
; PRIOR FILING DATE: 2001-11-27  
; PRIOR APPLICATION NUMBER: GB 0203185.4  
; PRIOR FILING DATE: 2002-02-11  
; NUMBER OF SEQ ID NOS: 306  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 142  
; LENGTH: 3081  
; TYPE: DNA  
; ORGANISM: Drosophila melanogaster  
; US-10-840-060-142

Query Match 100.0%; Score 3081; DB 9; Length 3081;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3081; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AAACATAAAATTTGTTGCTGACATCTGGTTCGCTTGCACAACTATTTCTAGCAGATTTT 60  
DB 1 AAACATAAAATTTGTTGCTGACATCTGGTTCGCTTGCACAACTATTTCTAGCAGATTTT 60  
QY 61 GTGATATTTCGTTGTCATCGGTTCGATTAATCCGCCAGTTTTTTTTTAAATGGAAGTGCT 120  
DB 61 GTGATATTTCGTTGTCATCGGTTCGATTAATCCGCCAGTTTTTTTTTAAATGGAAGTGCT 120  
QY 121 AACACATTTAGCGGTTCGGAAGATAGCAGGAAGACGCGGCTGCGGCTTTTCCTT 180  
DB 121 AACACATTTAGCGGTTCGGAAGATAGCAGGAAGACGCGGCTGCGGCTTTTCCTT 180

Sequence 1328, Ap  
Sequence 1328, Ap  
Sequence 486, App  
Sequence 5182, Ap  
Sequence 134, App  
Sequence 5, Appli  
Sequence 11, Appl  
Sequence 1386, Ap  
Sequence 1830, Ap  
Sequence 87, Appl  
Sequence 4857, Ap  
Sequence 146, App  
Sequence 294, App  
Sequence 294, App  
Sequence 293, App  
Sequence 293, App  
Sequence 1, Appli  
Sequence 3, Appli  
Sequence 7, Appli  
Sequence 7, Appli

181 TTGTGATATCGTGGCAGACGCAAGAAAACGACAGTTGGCATTTGAATTCAGACAAAC 240  
181 TTGTGATATCGTGGCAGACGCAAGAAAACGACAGTTGGCATTTGAATTCAGACAAAC 240  
241 ACACATCTAACGCCGACCGCAAGAGAGACACACACACACTCGGACACTCGAAAAAA 300  
241 ACACATCTAACGCCGACCGCAAGAGAGACACACACACACTCGGACACTCGAAAAAA 300  
301 AAAAACAAGACGCTGTCGGCGACCTCGACACAGCAGTTGGTTCGATTTAGTTGTCATGC 360  
301 AAAAACAAGACGCTGTCGGCGACCTCGACACAGCAGTTGGTTCGATTTAGTTGTCATGC 360  
361 CTTGAATTCGGTTCGGGCGTTAGTTTCCACAAGTTTATCGCTCGTCAAGAAAACAGGAAA 420  
361 CTTGAATTCGGTTCGGGCGTTAGTTTCCACAAGTTTATCGCTCGTCAAGAAAACAGGAAA 420  
421 TAAATTTATTTTCGACCTTAAATAATCTGACTAAATTTGTTTGTGTTTATGTTATTTAT 480  
421 TAAATTTATTTTCGACCTTAAATAATCTGACTAAATTTGTTTGTGTTTATGTTATTTAT 480  
481 TAGGCACATTTTGCACACACAAAGTACTTACTACATCTAGCATTAACGGAACTCCTCT 540  
481 TAGGCACATTTTGCACACACAAAGTACTTACTACATCTAGCATTAACGGAACTCCTCT 540  
541 GCAAGCAGTGGAGTTGCTGCTCAATCAAGCAGTACTCGGAGTTAACGAGGATAAGCCGG 600  
541 GCAAGCAGTGGAGTTGCTGCTCAATCAAGCAGTACTCGGAGTTAACGAGGATAAGCCGG 600  
601 GAGAAAGAAAGAGATCGGTGGAGATAGAGATATACAGTGGAGTCAAGAGAAAGGA 660  
601 GAGAAAGAAAGAGATCGGTGGAGATAGAGATATACAGTGGAGTCAAGAGAAAGGA 660  
661 TCATGGACATGATTACCGTGGGGGAGAGCGTCAAGATCAAGGGACGATCGCCGCTCC 720  
661 TCATGGACATGATTACCGTGGGGGAGAGCGTCAAGATCAAGGGACGATCGCCGCTCC 720  
721 ACATGGCCGTGGTGGCGTGTATCAACAGTGGGCAAGTGCATCACAGTCAATGGTACG 780  
721 ACATGGCCGTGGTGGCGTGTATCAACAGTGGGCAAGTGCATCACAGTCAATGGTACG 780  
781 AGCGCGGCAACGAAAGGCGCAAGAGGTAGAACTTGGAGCCATCTACGCTCAATCCGG 840  
781 AGCGCGGCAACGAAAGGCGCAAGAGGTAGAACTTGGAGCCATCTACGCTCAATCCGG 840  
841 AGCTAATCCAGATACTCTCGACAGCAGCGCCCGCCCGGAGCCCAAGAAACAAGCCACCG 900  
841 AGCTAATCCAGATACTCTCGACAGCAGCGCCCGCCCGGAGCCCAAGAAACAAGCCACCG 900  
901 CGCCGATGAACCTCTCGCGTAAATCCACACAATCGGCTATCGGTGGCAATCTCACACGCC 960  
901 CGCCGATGAACCTCTCGCGTAAATCCACACAATCGGCTATCGGTGGCAATCTCACACGCC 960  
961 GTATGACCAATGGCCGGAACATGTGAAACAAGATCCAGGAAAGCAGTTCGATTCCTCAATC 1020  
961 GTATGACCAATGGCCGGAACATGTGAAACAAGATCCAGGAAAGCAGTTCGATTCCTCAATC 1020  
1021 CGATTGTGACAGCAATAGCTGTGATACAAACAGCACTCCAACTACGCGCCGCGGAG 1080  
1021 CGATTGTGACAGCAATAGCTGTGATACAAACAGCACTCCAACTACGCGCCGCGGAG 1080  
1081 GTGTGGGACACAAACGTCGACGACCACTGGATTACAGCGTCCACGGTACTCGCAAGCTG 1140  
1081 GTGTGGGACACAAACGTCGACGACCACTGGATTACAGCGTCCACGGTACTCGCAAGCTG 1140  
1141 CTACCGGCGCAGAGCAGACAAAGATCGCTCGCGGTGCTTAATAACATTTGCCCAATC 1200  
1141 CTACCGGCGCAGAGCAGACAAAGATCGCTCGCGGTGCTTAATAACATTTGCCCAATC 1200  
1201 CCAGCGGCGCAGAGCAGTCTGCTCGCGGGCACAAGGAGTCCGCTCGGCGCACACCC 1260  
1201 CCAGCGGCGCAGAGCAGTCTGCTCGCGGGCACAAGGAGTCCGCTCGGCGCACACCC 1260  
1261 AGGGAGCTGGCGGCGTAGTATCCCGGGCGATCGCACGCAATTGAAAGAGGTGGAGGACTGA 1320

1261 AGGGAGCTGGCGGCGTAGTATCCCGGGCGATCGCACGCAATGAAAGAGGTGGAGGACTGA 1320  
1321 AGGAAATTCGCGAGAGCGCGCCCGACAGGCCGAGATGAAGAGAGAGGTGGCGC 1380  
1321 AGGAAATTCGCGAGAGCGCGCCCGACAGGCCGAGATGAAGAGAGAGGTGGCGC 1380  
1381 TGATGAACCAAGGATCCGGGCAATCCAACTGGGAGCGGCGCAATGATACGCAATATC 1440  
1381 TGATGAACCAAGGATCCGGGCAATCCAACTGGGAGCGGCGCAATGATACGCAATATC 1440  
1441 AGAGCAGCTGGAATTTGTCCGCTGCTCGATGCGCAGGCGGTGATGACCATCAGATCA 1500  
1441 AGAGCAGCTGGAATTTGTCCGCTGCTCGATGCGCAGGCGGTGATGACCATCAGATCA 1500  
1501 CAGTGTGCGTGGCGCAAGCGTCCCATTTAGCCGCAAGGAGTCAATCGCAAGGATCGATG 1560  
1501 CAGTGTGCGTGGCGCAAGCGTCCCATTTAGCCGCAAGGAGTCAATCGCAAGGATCGATG 1560  
1561 TCATTTTCGTTGCGCGCAAGGACATGCTCATCGTGACGAGCGCGCAGCAAGGTGACCC 1620  
1561 TCATTTTCGTTGCGCGCAAGGACATGCTCATCGTGACGAGCGCGCAGCAAGGTGACCC 1620  
1621 TCACCAAGTTCTCGGAGAACCAAGTTTTCGCTTCGACTAGCCCTTCAACGACACGTGCG 1680  
1621 TCACCAAGTTCTCGGAGAACCAAGTTTTCGCTTCGACTAGCCCTTCAACGACACGTGCG 1680  
1681 ACAATGCGATGGTATACAAATACACAGCAAGCGTTGGTGAACCAATTTTCAGGGCG 1740  
1681 ACAATGCGATGGTATACAAATACACAGCAAGCGTTGGTGAACCAATTTTCAGGGCG 1740  
1741 GAATGGCGAGTCTTCGCTTACGCGCAGAGGATCGGCGCAACCAACCAACCATGGCG 1800  
1741 GAATGGCGAGTCTTCGCTTACGCGCAGAGGATCGGCGCAACCAACCAACCATGGCG 1800  
1801 GTGAGTTTAAATGGAAGGTGCGAGACTGCAAGAACGGCATCTAGCCATGCGCGCCAAAG 1860  
1801 GTGAGTTTAAATGGAAGGTGCGAGACTGCAAGAACGGCATCTAGCCATGCGCGCCAAAG 1860  
1861 ATGTCTTTTGTGACCTGAAATATGCGCGTTTACCGCGCATGAATCTAGTCTCGGCCA 1920  
1861 ATGTCTTTTGTGACCTGAAATATGCGCGTTTACCGCGCATGAATCTAGTCTCGGCCA 1920  
1921 GTTCTCTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTCTGTCGCAACGACGAACTGC 1980  
1921 GTTCTCTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTCTGTCGCAACGACGAACTGC 1980  
1981 GCGTCTCTGAGGATGTTAAACAGCAAGTGCAGGTGGTGGGACTCACCGAGAGGTGGTGC 2040  
1981 GCGTCTCTGAGGATGTTAAACAGCAAGTGCAGGTGGTGGGACTCACCGAGAGGTGGTGC 2040  
2041 ATGCGCTGAGAGGATGTAAGTCAATCAGACGCGCAATGTGCGCGCAACATCCGCGCC 2100  
2041 ATGCGCTGAGAGGATGTAAGTCAATCAGACGCGCAATGTGCGCGCAACATCCGCGCC 2100  
2101 AGAGCTCGGCCCACTCAATTCGTCGCTTCGCAACGCGCTTTCCAGATTTGCTGCGGC 2160  
2101 AGAGCTCGGCCCACTCAATTCGTCGCTTCGCAACGCGCTTTCCAGATTTGCTGCGGC 2160  
2161 CGCAGGCTCGAGCAAGATCCATGGCAAGTTCTCGTTTCATCGATCTCGCGCGCAATGAGC 2220  
2161 CGCAGGCTCGAGCAAGATCCATGGCAAGTTCTCGTTTCATCGATCTCGCGCGCAATGAGC 2220  
2221 GGGCGTGGACACTTCTCTCGCGCGATCGGCAGACGCGTATGAGGGTGGCGAGATTAAAC 2280  
2221 GGGCGTGGACACTTCTCTCGCGCGATCGGCAGACGCGTATGAGGGTGGCGAGATTAAAC 2280  
2281 AATCGCTGCTGGCCCTCAAGGAGTGCATTCGCTGGTGGGCAACACGCTCGGCCCACTTGC 2340  
2281 AATCGCTGCTGGCCCTCAAGGAGTGCATTCGCTGGTGGGCAACACGCTCGGCCCACTTGC 2340  
2341 CCTTCGCTGCTCCAAACTCACCGAGTGTGCGCGACTCGTTTCATTTGGGCGAGAGGACA 2400  
2341 CCTTCGCTGCTCCAAACTCACCGAGTGTGCGCGACTCGTTTCATTTGGGCGAGAGGACA 2400

```
Db 2341 CCTTCGGTGTCTCCAAACTCACCAGGTGCTCGCGACTCGTTCATTGGCGGAGAGCA 2400
Qy 2401 AGACGTGCATGATAGCATGATCTCCCGGACTTAGCTCTCGCGAGCACAGCTCAACA 2460
Db 2401 AGACGTGCATGATAGCATGATCTCCCGGACTTAGCTCTCGCGAGCACAGCTCAACA 2460
Qy 2461 CGCTGGCTATCGGATCGTGTCAAGAGCTGTGTGCTCAAGGATATCGTCAAGTTTGGC 2520
Db 2461 CGCTGGCTATCGGATCGTGTCAAGAGCTGTGTGCTCAAGGATATCGTCAAGTTTGGC 2520
Qy 2521 CTGGCGGCGACACCGAGGCCATCGAGATCAGCGACGACGAGGAGGAGGAGCTCAACA 2580
Db 2521 CTGGCGGCGACACCGAGGCCATCGAGATCAGCGACGACGAGGAGGAGGAGCTCAACA 2580
Qy 2581 TGTGTGATCGCATCTCGATCTAGCTGATCCCAATTCGATGACCGCGGAGGAGTGA 2640
Db 2581 TGTGTGATCGCATCTCGATCTAGCTGATCCCAATTCGATGACCGCGGAGGAGTGA 2640
Qy 2641 ATAATCAGCGTGTCTCGGCTCTCATCATCTCGGGGGGCTCATTCACCAATTAATA 2700
Db 2641 ATAATCAGCGTGTCTCGGCTCTCATCATCTCGGGGGGCTCATTCACCAATTAATA 2700
Qy 2701 ACAACAACAAGAACGGAACCGCGCAACATGAGCTGCGCCATGCTGAGTTCGCTGAGCG 2760
Db 2701 ACAACAACAAGAACGGAACCGCGCAACATGAGCTGCGCCATGCTGAGTTCGCTGAGCG 2760
Qy 2761 AACACGAGATGTCGACGAGCTGATTTGTCAGCACACAGGCCATCGACGACCTGCGAGCAGA 2820
Db 2761 AACACGAGATGTCGACGAGCTGATTTGTCAGCACACAGGCCATCGACGACCTGCGAGCAGA 2820
Qy 2821 CGGAGGAGATGTTGGTGGAGTATCATCGCACCGCTTAATGCCACACTGGAGACTTCCTCG 2880
Db 2821 CGGAGGAGATGTTGGTGGAGTATCATCGCACCGCTTAATGCCACACTGGAGACTTCCTCG 2880
Qy 2881 CCGAGTCGAAGCGCTGTACAATCTGACCAACTATGTGACTACGACGAGGACTGCTACT 2940
Db 2881 CCGAGTCGAAGCGCTGTACAATCTGACCAACTATGTGACTACGACGAGGACTGCTACT 2940
Qy 2941 GCAAAACGGGCGAGTCGATGTTCTCGACGCTGTCGACATCGCCATCCAGTCCCGGACGA 3000
Db 2941 GCAAAACGGGCGAGTCGATGTTCTCGACGCTGTCGACATCGCCATCCAGTCCCGGACGA 3000
Qy 3001 TGATGGCCGAATATCGCGCAAGTTGGCCAAAGAGAGATGCTGTCGTCAGCTTCAATT 3060
Db 3001 TGATGGCCGAATATCGCGCAAGTTGGCCAAAGAGAGATGCTGTCGTCAGCTTCAATT 3060
Qy 3061 CGCCGAATGGCAAGCGTTAGT 3081
Db 3061 CGCCGAATGGCAAGCGTTAGT 3081
```

RESULT 2

```
US-11-097-143-1103
; Sequence 1103, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION: DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
```

```
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1103
; LENGTH: 3081
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-1103

Query Match 100.0%; Score 3081; DB 10; Length 3081;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3081; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AAACATAAAAAATTGTGTCTGACATCTGGTGGCTTGCAGAACTATTTCTAGCAGATTTT 60
Db 1 AAACATAAAAAATTGTGTCTGACATCTGGTGGCTTGCAGAACTATTTCTAGCAGATTTT 60
Qy 61 GTGATATTTTGTGTGTGATCGGTGATATAATCCGCCAGTTTTTTTTTAATGGAAGTGT 120
Db 61 GTGATATTTTGTGTGTGATCGGTGATATAATCCGCCAGTTTTTTTTTAATGGAAGTGT 120
Qy 121 AACACATTTGTAGCGGTTCGGAAGATAGCAGGAAGAGCAGCGGCTGCGTTTTTCCCTT 180
Db 121 AACACATTTGTAGCGGTTCGGAAGATAGCAGGAAGAGCAGCGGCTGCGTTTTTCCCTT 180
Qy 181 TTTGTTATCCGTTGTCAGACGCAACGAAACGACAGTTTGGCAATTTGAATTCAGCACAAC 240
Db 181 TTTGTTATCCGTTGTCAGACGCAACGAAACGACAGTTTGGCAATTTGAATTCAGCACAAC 240
Qy 241 ACACATATTAACCGCGACCCGCAAGCAGCACACACACACTCGGACACTCGAAAAAA 300
Db 241 ACACATATTAACCGCGACCCGCAAGCAGCACACACACACTCGGACACTCGAAAAAA 300
Qy 301 AAAAAACAGACGCTGTCGCGGACCTCGCAAGCAGTTCGGTTTCGATTTAGTTGTCAATGC 360
Db 301 AAAAAACAGACGCTGTCGCGGACCTCGCAAGCAGTTCGGTTTCGATTTAGTTGTCAATGC 360
Qy 361 CTTGAATTCGGTTCGGGCTTAGTTTCCACAAAGTTTATCGCTCGTCAAGAAACAGAA 420
Db 361 CTTGAATTCGGTTCGGGCTTAGTTTCCACAAAGTTTATCGCTCGTCAAGAAACAGAA 420
Qy 421 TAAATTTATTTTCGACCTTAAAAATCTGACTTAAATTTGTTTTTTTGTATTTATTTATT 480
Db 421 TAAATTTATTTTCGACCTTAAAAATCTGACTTAAATTTGTTTTTTTGTATTTATTTATT 480
Qy 481 TAGGCACATTTTGCACACCAACGCTAGTTACTACTACATCTAGCAGTAAACGGAACCTCT 540
Db 481 TAGGCACATTTTGCACACCAACGCTAGTTACTACTACTACTAGCAGTAAACGGAACCTCT 540
Qy 541 GCAAGCAGTGAAGTTGCTGCTCCATCAAGCAGTACTCGGAGTTAACGCGAGTAAAGCGG 600
Db 541 GCAAGCAGTGAAGTTGCTGCTCCATCAAGCAGTACTCGGAGTTAACGCGAGTAAAGCGG 600
Qy 601 GAGAAAGAAAGAGATCGGTGGAGATATAGAGATATACAGTGGAGTCAAGAGAGAGGA 660
Db 601 GAGAAAGAAAGAGATCGGTGGAGATATAGAGATATACAGTGGAGTCAAGAGAGAGGA 660
Qy 661 TCATGGACATGATTTACGGTGGGCGAGCGTCAAGATCAAGCGGAGTGGCCGCGTCC 720
Db 661 TCATGGACATGATTTACGGTGGGCGAGCGTCAAGATCAAGCGGAGTGGCCGCGTCC 720
Qy 721 ACATGGCGGTGGTGGCTGATCAACAGTCCGGCAAGTGCATCACAGTCAATGTTAGT 780
Db 721 ACATGGCGGTGGTGGCTGATCAACAGTCCGGCAAGTGCATCACAGTCAATGTTAGT 780
Qy 781 AGCGCGGCAACGAAAGGCGAGGAGTGAAGTGGAGCCATCTACGCTCAATCCGG 840
Db 781 AGCGCGGCAACGAAAGGCGAGGAGTGAAGTGGAGCCATCTACGCTCAATCCGG 840
```

Db 781 AGCGCGCGAAGCAAGGAGGAGTAGAATCGAGCCATCTCACGCTCAATCCGG 840  
Qy 841 AGCTAATGCAAGATACTGTGCAACAGCAGCGCCCGGAGCCCAAGCAACGCG 900  
Db 841 AGCTAATGCAAGATACTGTGCAACAGCAGCGCCCGGAGCCCAAGCAACGCG 900  
Qy 901 CGCCGATGAACCTCTCGGTAATCCACACAAATCGGCTATCGGTGGCAATCTCACAGCC 960  
Db 901 CGCCGATGAACCTCTCGGTAATCCACACAAATCGGCTATCGGTGGCAATCTCACAGCC 960  
Qy 961 GTATCACCATGGCCGGAACATGCTGAACAGATCCAGAAAGCCAGTCTGATTCCTCAATC 1020  
Db 961 GTATCACCATGGCCGGAACATGCTGAACAGATCCAGAAAGCCAGTCTGATTCCTCAATC 1020  
Qy 1021 CCAATTGTCAAGCAGCAATAGCTGAATACAAACAGCAACTCCAACTACAGCCCGCGGAG 1080  
Db 1021 CCAATTGTCAAGCAGCAATAGCTGAATACAAACAGCAACTCCAACTACAGCCCGCGGAG 1080  
Qy 1081 GTGTGGGCAACCAACGTCGACGACCACTGGATTAACAGCTGCTGATTCGCAAGCTG 1140  
Db 1081 GTGTGGGCAACCAACGTCGACGACCACTGGATTAACAGCTGCTGATTCGCAAGCTG 1140  
Qy 1141 CTACCGGCGAGCAGCAGACAGGATCGCTCGCGGTGCTTAATAACATGTCGCAATC 1200  
Db 1141 CTACCGGCGAGCAGCAGACAGGATCGCTCGCGGTGCTTAATAACATGTCGCAATC 1200  
Qy 1201 CCAGCGCGCAGCAGCTGCTGTCGCGGCGACAAAGAGTCCCACTGCGGCCCAACCC 1260  
Db 1201 CCAGCGCGCAGCAGCTGCTGTCGCGGCGACAAAGAGTCCCACTGCGGCCCAACCC 1260  
Qy 1261 AGGAGCTGGCGGCGCTAGTACCGCGGATCGCAGCAATTGAAGAGGTGGAGCGACTGA 1320  
Db 1261 AGGAGCTGGCGGCGCTAGTACCGCGGATCGCAGCAATTGAAGAGGTGGAGCGACTGA 1320  
Qy 1321 AGGAGATCGGAGAGCGAGCGCGCCGACAGGCGCGAGATGAAGAGAGGTGGCGC 1380  
Db 1321 AGGAGATCGGAGAGCGAGCGCGCCGACAGGCGCGAGATGAAGAGAGGTGGCGC 1380  
Qy 1381 TGATGAACCAAGATCCGCGCAATCCAACTCGGAGAGCGCGCAATGATCGCAATATC 1440  
Db 1381 TGATGAACCAAGATCCGCGCAATCCAACTCGGAGAGCGCGCAATGATCGCAATATC 1440  
Qy 1441 AGAGCAGCTGGAAATTTGTGCGCTGCTCGATGCGCAGCGCGTGGATGACCATCAGATCA 1500  
Db 1441 AGAGCAGCTGGAAATTTGTGCGCTGCTCGATGCGCAGCGCGTGGATGACCATCAGATCA 1500  
Qy 1501 CAGTGTGCTGCGCAAGCGTCCATTTAGCCCAAGAGGTCAATCGCAAGAGATCGATG 1560  
Db 1501 CAGTGTGCTGCGCAAGCGTCCATTTAGCCCAAGAGGTCAATCGCAAGAGATCGATG 1560  
Qy 1561 TCATTTGCTGCGCGCAAGCAGATGCTTCATGTCGACGAGCGCGCAGCAAGGTGCGACC 1620  
Db 1561 TCATTTGCTGCGCGCAAGCAGATGCTTCATGTCGACGAGCGCGCAGCAAGGTGCGACC 1620  
Qy 1621 TCACCAAGTTCCTGGAGAACCAAGTTTCGCTTCGACTAGCCCTTCAACGACACGTGCG 1680  
Db 1621 TCACCAAGTTCCTGGAGAACCAAGTTTCGCTTCGACTAGCCCTTCAACGACACGTGCG 1680  
Qy 1681 ACAATGCCATGGTATACAAATACACAGCAAGCGTTGGTGAAACCATTTCAGGGCG 1740  
Db 1681 ACAATGCCATGGTATACAAATACACAGCAAGCGTTGGTGAAACCATTTCAGGGCG 1740  
Qy 1741 GAATGGGAGCTGCTCGCTTACGCGCAGACGGATCGGGCAAAACCCACCATGGGCG 1800  
Db 1741 GAATGGGAGCTGCTCGCTTACGCGCAGACGGATCGGGCAAAACCCACCATGGGCG 1800  
Qy 1801 GTGAGTTTAATGGAAGAGTGCAGGACTGCAAGAACGGCATCTAGCCATGCGCGCCAAAG 1860  
Db 1801 GTGAGTTTAATGGAAGAGTGCAGGACTGCAAGAACGGCATCTAGCCATGCGCGCCAAAG 1860  
Qy 1861 ATGCTTTGTGACCTGAAATGCGGTTTACCGCGCATGAATCTAGTCTCGGCCA 1920  
Db 1861 ATGCTTTGTGACCTGAAATGCGGTTTACCGCGCATGAATCTAGTCTCGGCCA 1920

Qy 1921 GTTCTTTGAGATTTACAGTGGGCAAGCTCTTCGATCTTCTGTCGACAGCAGAAACTGC 1980  
Db 1921 GTTCTTTGAGATTTACAGTGGGCAAGCTCTTCGATCTTCTGTCGACAGCAGAAACTGC 1980  
Qy 1981 GCGTCTTGGAGGATGGTAAACAGCAAGTGCAGGTGGGACTCACCGAAGGTGGTCG 2040  
Db 1981 GCGTCTTGGAGGATGGTAAACAGCAAGTGCAGGTGGGACTCACCGAAGGTGGTCG 2040  
Qy 2041 ATGCGCTGAGAGGATCTGAAGCTCATCCAGCACGGCAATGTGCCCAACATCCGGCC 2100  
Db 2041 ATGCGCTGAGAGGATCTGAAGCTCATCCAGCACGGCAATGTGCCCAACATCCGGCC 2100  
Qy 2101 AGAGCTCGGCGCACTCTCAATTCGTCGCTTCGCGACCGGCTTTCAGATGTGTCGGCC 2160  
Db 2101 AGAGCTCGGCGCACTCTCAATTCGTCGCTTCGCGACCGGCTTTCAGATGTGTCGGCC 2160  
Qy 2161 CGCAGGCTCGAGCAAGATCCATGGCAAGTTCTCGTTTCATCGATCTCGCGGCAATCAGC 2220  
Db 2161 CGCAGGCTCGAGCAAGATCCATGGCAAGTTCTCGTTTCATCGATCTCGCGGCAATCAGC 2220  
Qy 2221 GGGCGTGGACACTTCTCGGCGGATCGGCAGACCGCTATGAGGGGTGCGGAGATTAAAC 2280  
Db 2221 GGGCGTGGACACTTCTCGGCGGATCGGCAGACCGCTATGAGGGGTGCGGAGATTAAAC 2280  
Qy 2281 AATCGCTGCTGGCCCTCAAGAGTGCATTCGTCGTTGGGCAACAGTCCGCCACTTGC 2340  
Db 2281 AATCGCTGCTGGCCCTCAAGAGTGCATTCGTCGTTGGGCAACAGTCCGCCACTTGC 2340  
Qy 2341 CCTTCGCTGCTCCAACTCACCCAGGTGCTGGCGGACTCGTTTCATTGGCGAGAGAGCA 2400  
Db 2341 CCTTCGCTGCTCCAACTCACCCAGGTGCTGGCGGACTCGTTTCATTGGCGAGAGAGCA 2400  
Qy 2401 AGAGCTGCATGATAGCATGATCTCGCGGGAATTAGCTCTCGAGCACAACGCTCAACA 2460  
Db 2401 AGAGCTGCATGATAGCATGATCTCGCGGGAATTAGCTCTCGAGCACAACGCTCAACA 2460  
Qy 2461 CGCTGCTATGCGGATCGTCAAGAGCTGCTGCTCAAGGATATGTCGTAAGTTTGC 2520  
Db 2461 CGCTGCTATGCGGATCGTCAAGAGCTGCTGCTCAAGGATATGTCGTAAGTTTGC 2520  
Qy 2521 CTGCGCGCAGACCGAGCCCATCGAGATCACCGACGACGAGGAGGAGGCTCAACA 2580  
Db 2521 CTGCGCGCAGACCGAGCCCATCGAGATCACCGACGACGAGGAGGAGGCTCAACA 2580  
Qy 2581 TGGTGCATCCGCACTGCGATGCAATCCCAATTCGCAATGCAACGCGCCAGCGAGTCGA 2640  
Db 2581 TGGTGCATCCGCACTGCGATGCAATCCCAATTCGCAATGCAACGCGCCAGCGAGTCGA 2640  
Qy 2641 ATAATCAGCGTGTCCGGCTCTCATCTCGGGGGGGTCAATCAACAATAAATA 2700  
Db 2641 ATAATCAGCGTGTCCGGCTCTCATCTCGGGGGGGTCAATCAACAATAAATA 2700  
Qy 2701 ACAACCAAGAACCGGAAACCGCGCAACATGAGACCTGCGCCATGCTGAGTCTGAGCG 2760  
Db 2701 ACAACCAAGAACCGGAAACCGCGCAACATGAGACCTGCGCCATGCTGAGTCTGAGCG 2760  
Qy 2761 AACACAGATGTCGAGAGCTGATTTGTCAGACACAGCGCATCGAGACTCGCAGCAGA 2820  
Db 2761 AACACAGATGTCGAGAGCTGATTTGTCAGACACAGCGCATCGAGACTCGCAGCAGA 2820  
Qy 2821 CGCAGAGATGTTGGTGGAGTATCATCGCACCGTTAATGCCACACTGGAGACCTTCTCG 2880  
Db 2821 CGCAGAGATGTTGGTGGAGTATCATCGCACCGTTAATGCCACACTGGAGACCTTCTCG 2880  
Qy 2881 CCGAGTTCGAGGCGCTGTACAATCTGACCAACTATGTGGACTACGACAGGACTCGTACT 2940  
Db 2881 CCGAGTTCGAGGCGCTGTACAATCTGACCAACTATGTGGACTACGACAGGACTCGTACT 2940  
Qy 2941 GCAAAACGGGCGAGTGCATGTTCTCGCAGTGTGCACTGCGCATCTGAGTCCGCGGACA 3000  
Db 2941 GCAAAACGGGCGAGTGCATGTTCTCGCAGTGTGCACTGCGCATCTGAGTCCGCGGACA 3000









```
; SEQ ID NO 5990
; LENGTH: 1970
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-5990

Query Match      16.4%; Score 505.8; DB 10; Length 1970;
Best Local Similarity 64.7%; Pred. No. 4.3e-140;
Matches 822; Conservative 0; Mismatches 432; Indels 17; Gaps 4;

QY 1252 CCACAAACCAGGAGTGGCGCGGTAGTACCCGGCGATCGCACGCAATGAAAGAGGTGG 1311
DB 399 CGCGCACGCGCAGACGCGCGCGCGTTCGAGAGGA--CGAGGTGGTGCACACAGGCTG 456
QY 1312 ACGGACTGAAGCAGAAATCCGAGAGCGACGCGCCCGCAGCGCCGAGATGAAGGAGGAGA 1371
DB 457 AAAGAAATGCGAAAGAACCGGAACGAGCGGAGAGACGACAAACCCAGGACTCGTCTAGATC 516
QY 1372 AGGTGGCGCTGATGAACCAAGGATCCGGGCAATCCAAACTGGGAGACGGCGCAAAATGATAC 1431
DB 517 GGGAGCAGGGGAAGAACGAGATCCGGGAATCCCAACTGGGAAGTAGCCAGAAATGATAC 576
QY 1432 GCGAATATCAGACGAGCTGGAAATTTGTGCCCTGCTCGATGG-----CCAGGCGG 1482
DB 577 GACTGCAACGCGAGCAAAATGGAGAGTCAAGCGGTGAGAAAGTGTACTACGAACGAACGAA 636
QY 1483 TCGATGACCATCAGATCAGATGTCGTCGCGCAAGCGTCCCATTTAGCCGCAAGAGGTCA 1542
DB 637 TCAATTTGCCACCAATATATGTTTGTGTGAGGAAGAGACCACTGAGGCGCAAGGAGTGG 696
QY 1543 ATGCCAAGGAGATCGATTCATTTCCGTCGCCGCAAGGACATGCTCATCTGTCACGAGC 1602
DB 697 CTGACCGGGAACAGGATGTGTGTCAGCATTCGCTCTAAGCACACATTTGTTGGTCCACGAGC 756
QY 1603 CGCGCAGCAGAGTGCAGCTCACCAGTTCTCTGGAGAACACAAAGTTTCGTTTGCATGACG 1662
DB 757 CCGCGAAGCATGTGAACCTGTGTCAAGTTCTCTGGAATAATCATAGCTTCGCTTTCGATTACG 816
QY 1663 CTTTCAACGACACGTCGACATGTCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1722
DB 817 TCTTCGACGAGAGTGCTTCAATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 876
QY 1723 AAACCATTTTCGAGGCGGAAATGGCGAGTGTCTTCGCTTACGGCCAGACGGGATCGGCA 1782
DB 877 AGCACAATTTTGATGCGGAAATGGCCAGTGTTCGCTTACGGACAAACTGGAACGCGCA 936
QY 1783 AAACCCACACATGGGCGGTGAGTTTATGAAAGGTGAGGACATGCAAGAACGCGCATCT 1842
DB 937 AGACCTATACGATGGGTGTGTCAGTTCCCGGAAAGGCATCAGAGCTCAATGGATGGCATCT 996
QY 1843 ACGCCATGGCGGCAAGGATGTCCTTGTGACCCCTCAATATCCGGTTTACCGGCGCATGA 1902
DB 997 ATGCAATGGCCGCTAAGGACGTTTCTCCACTCTAAGACGTTTCCCTATACCAAGCTTA 1056
QY 1903 ATCTAGTGTCTCGGCCAGATTTCTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTC--- 1959
DB 1057 ATCTGAAAGTTTACTGCGAGCTTCTTCGAGATCTACGCGCACCGGCTGTCGATCTCTCTGA 1116
QY 1960 TGTCCGACACGAGAAATCGCGCTCTGGAGGATGGTAAACAGCAAGTGCAGAGTGGTGG 2019
DB 1117 TGCCTGGGAAGCCACAACTCGCTGTCTTTGGAGGATAGAAACACAGCAGGTGCAAGTGGTGG 1176
QY 2020 GACTCACCGAAGAGTGTTCGATGCGTCGAGGAGTACTGAAGCTCATCCAGCACGCGCA 2079
DB 1177 GCCTCACCGAGATTCAGTACGAACACCGCGGAAGTTCTGACCTACTCAGATGGGCA 1236
QY 2080 ATGTGCGCGCAACATCCGCGCAGACGTCCGCCCACTCCAATTTGCTGCGGTTTCGACGCGG 2139
DB 1237 ATAGTGTCCGAACCTCGGCTCACACCTCTGCGCAATTCCAAGTCTCTCCGATCGCATGCTG 1296
QY 2140 TTTTCCAGATGTGCTGCGCGCGCAGGCTCGACGAAGATCGATGGCAAGTTCTCGTTCA 2199
DB 1297 TGTTCCAATCGTGTGAGATCCGCGCGCGGCGAGAGCTACACGGGAAATTTCTCGCTTA 1356

RESULT 7
US-11-097-143-5989/c
; Sequence 5989, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5989
; LENGTH: 3970
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-5989

Query Match      16.4%; Score 505.8; DB 10; Length 3970;
Best Local Similarity 64.7%; Pred. No. 6.5e-140;
Matches 822; Conservative 0; Mismatches 432; Indels 17; Gaps 4;

QY 1252 CCACAAACCAGGAGTGGCGCGCTAGTACCCGGCGATCGCACGCAATGAAAGAGGTGG 1311
DB 2572 CCGCACCGCAGCAGCGCGCGCGCTAGTTCGAGAGGA--CGAGGTGGTGCACCAAGGCTG 2515
QY 1312 AGCGACTGAAGGAGAAATCGGAGAGGAGCGCGCCCGCAGCGCGGATGAAGGAGGAGA 1371
```

Db 2514 AAGAAATGCCAAAGAACGCGGACGAGGCGAGAGACGACAAAGCCAGGACTCGTCTAGATC 2455  
Qy 1372 AGTGGCGCTGATGAACACAGGATCCGGGCAATCCAACTGGGAGACGGCGCAATATGATAC 1431  
Db 2454 GGGAGCAGGGGAAAGAAAGAAAGATCCGGGAAATCCCAACTGGGAAGTAGCCAGAATGATAC 2395  
Qy 1432 GCGAATATCAGACGACGCTCGAATTTGTGCGGCTGCTCGATGG-----CCAGGCCG 1482  
Db 2394 GACTGCMACCGGAGCAAAATGGAGAGTCAGCGGGTGAGAAGTGGTACTACGAACGAACGAA 2335  
Qy 1483 TCGATGACCATCAGATCACAAGTGTGGCTGCGCAAGGCTCCATTAGCGCGCAAGGAGTCA 1542  
Db 2334 TCAATTGCCACCAATATATGTTGTGTGAGGAAGAGACCACTAGCGCGCAAGGAGCTGG 2275  
Qy 1543 ATCGAAGAGATCGATGTCATTTTCGGTCCGCGCAAGGACATGCTCATGTCACAGAGC 1602  
Db 2274 CTGACCGGGAAACAGGATGTGGTCAGCATTCGGTCTAAGCACACATTTGGTGGTCCACGAGC 2215  
Qy 1603 CGCGCAGCAGGTGCGACCTCACCAGTTCTTGAGAACCAAGTTTCGTTTCGACTACG 1662  
Db 2214 CCGCAAGCATGTGAACCTGGTCAAGTTCTTGGAATCATAGCTTCCGTTTCGATTACG 2155  
Qy 1663 CTTTCAACGACAGCTGCGCAATGCCATGGTATACAAATACACAGCAAGCCGTTGCTGA 1722  
Db 2154 TCTTCGACGAGGATGCTCCAATGCCACGGTCTACGAATTCACAGCCGACCCCTTGATTA 2095  
Qy 1723 AAACCAATTTTCGAGGCGGAATGGCAGCTGCTTCGCTACGCGCAGACGGGATCGGCA 1782  
Db 2094 AGCACATTTTGATGCGGAATGGCCAGCTGTTTCGCCCTAGCGACAACTGGAAGCGGCA 2035  
Qy 1783 AAACGACACCATGGGCGGTGAGTTTANTGGAAGGTGAGACTCGAGACTCAAGACGGCATCT 1842  
Db 2034 AGACCTATACGATGGGTGGTCAAGTTCCCGGAGGATCAGAGCTCAATGGATGGCATCT 1975  
Qy 1843 ACGCCATGGCGGCCAAGGATGCTTTGTGACCTGAATATGCGCGGTACCGCGCCATCA 1902  
Db 1974 ATGCAATGGCGCTAAGGACGTTTCTCCACTTAAGACGGTTCCCTATTAACAGCTTA 1915  
Qy 1903 ATCTAGTCTGCTCGGCCAGTTCTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTC--- 1959  
Db 1914 ATCTGAAAGTTTACTCGAGCTTCTTCGAGATCTACGCGCACCCGGGTTTCGATCTCCTGA 1855  
Qy 1960 TGTCCGACAGCAGAACTGCGGCTCTGAGAGATGGTAAACGAGAGTGCAGGTGGTGG 2019  
Db 1854 TGCTGGCAGGCACAACTGCGTGTCTTGAGGATAGAAACAGCAGGTGCAAGTGGTGG 1795  
Qy 2020 GACTCACCGAGAGGTGTCGATGGGTCGAGAGGTACTGAAGCTCATCCAGCAGGCA 2079  
Db 1794 GCCTCACCGAGATCCAGTACAGAACACCGCGGAGTCTGGACCTACTCGAGTTGGGCA 1735  
Qy 2080 ATGCTGCCGGAACATCCGGCCAGACGTCGGCCAACTCCAAATTCGTGCGCTTCGACGCGG 2139  
Db 1734 ATAGTCTCCGAACTCGGGTCACACCTCTGCCAAATTCGAAGTCTCCGATCGCATGCTG 1675  
Qy 2140 TTTTCCAGATTTGCTGCGGCGCAGGGCTCGACGAAGATCCATGGCAAGTTCTCGTTCA 2199  
Db 1674 TGTTCAAATCGTGTGAGATCCGCGCGGCGGAGAGCTACACGGGAAATTCCTCGCTTA 1615  
Qy 2200 TCGATCTGGCGGCAATGAGCGGGCGTGACACTTCTCGCCGATCCGCGAGACCGGTA 2259  
Db 1614 TAGATCTGGCGGGAATGAAAGAGGAGCGGACAAACAGCTCGCGGATCCAGACGCGCC 1555  
Qy 2260 TGGAGGTCCGAGATTAAACAAATCGCTGCGCCCTCAAGAGTGCATTTCTGTGGTTGG 2319  
Db 1554 TGGAGGATCCGAGATCAATNAATCGTCTGGTCTCTCAAGGATGCAATTCGGCTCTGG 1495  
Qy 2320 GMAACAGTCCGCCACTTGGCCCTTCGGTGTCTCCAACTCACCCAGGTGCTGGCGACT 2379  
Db 1494 GCGCGCAGTCTGATCAATTTGGCAATTCGGTGGTCTCAAGCTGACCCAAAGTCTCTCGCGGACT 1435  
Qy 2380 CGTTCAATTCGCG---AGAGAGCAGACGTCGATGATGACCATGATCTCGCGGACTTA 2436  
Db 1434 CTTTATCGGAGGTAAAGAGGTGAAACCTGTCATGATGCCATGATCTCGCCATGCTTGC 1375

Qy 2437 GCTCTCGAGCACACGCTCAACACGCTGCGCTATCGCGATCGTGTCAAGGAGCTGGTGG 2496  
Db 1374 ATTGCGTGGAGCATACCTTTGAACAGCTCGGTTATCGCGATCGGTTGAAGGAACTAAGTG 1315  
Qy 2497 TCNAGGATATC 2507  
Db 1314 TGGAGTCGATC 1304

RESULT 8

US-10-651-510-1  
; Sequence 1, Application US/10651510  
; Publication No. US20040096949A1  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1042A  
; CURRENT APPLICATION NUMBER: US/10/651,510  
; PRIOR FILING DATE: 2003-08-28  
; PRIOR APPLICATION NUMBER: US/09/724,215  
; PRIOR FILING DATE: 2008-11-28  
; PRIOR APPLICATION NUMBER: 09/295,612  
; PRIOR FILING DATE: 1999-04-20  
; PRIOR APPLICATION NUMBER: 09/314,464  
; PRIOR FILING DATE: 1999-05-18  
; PRIOR APPLICATION NUMBER: 09/594,669  
; PRIOR FILING DATE: 2000-06-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1292  
; TYPE: DNA  
; ORGANISM: Human  
US-10-651-510-1

Query Match 14.8%; Score 454.6; DB 7; Length 1292;  
Best Local Similarity 62.6%; Pred. No. 7.9e-125;  
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;

Qy 1300 TCAGAGAGGTGGAGGACTGAAGAGAAATCGCGAAGCGACGCGCCGACAGGCGGAGA 1359  
Db 22 TGAAGGAAGTGGAAAAAATGAAGAACAGCGAGAGAGAAAGGCCAGAACTCTGAAA 81  
Qy 1360 TGAAGGAGGAGAGGTGGCGCTGTGAACACGAGATCCGGGCAATCCAAACTGGGACCG 1419  
Db 82 TGAGAAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTATTTCCAACTGGGAATTTG 135  
Qy 1420 CGCAAAATGATACGCGAAATATCAGACGACGCTGCGAATTTGTGCGGCTGCTCGATGGCCAG 1479  
Db 136 CCGGAATGATTAAGAAATTTGGGGTACTTTGGGAATGTCATCCACTTACTATGACTGATC 195  
Qy 1480 CCGTGCATGACCATCAGATCAGTGTGCGTGCAGAGCGTCCCATTTAGCCGCAAGGAGG 1539  
Db 196 CTATCGAAGAGCACAGAATATGTGCTGTGTTAGGAAACGCCCACTGAATAAGCAAGAAAT 255  
Qy 1540 TCAATCGCAGGAGATCGATGTCATTTTCGGTGGCCGCGCAGAGACATGCTCATCGTGACG 1599  
Db 256 TGGCCAAGAAAATGATGTAATTTCCATTTCTTAGCAAGTGTCTCCTCTTTGGTACATG 315  
Qy 1600 AGCCGCGCAGCAAGGTCCACCTCACCAAGTTCTCGAGAGAACCAAAAGTTTCGCTTCGACT 1659  
Db 316 AACCCAAAGTTGAAAGTGAGCTTAACAAAGTATCTCGAGAACCAAGCATTCGCTTTGACT 375  
Qy 1660 AGCCTTCAACGACAGCTGCGCAATGCCATGGTATACAAATACACAGCCAGCCGTTGG 1719  
Db 376 TTGCATTTGATGAACACAGCTTCGAATGAATGTTGTCTACAGGTTTCACAGAGGCCACTGG 435  
Qy 1720 TGAACCACTTTTCGAGGCGGGAATGGCGAGCTTTCGCTTACGGCCAGACGCGGATCGG 1779  
Db 436 TACAGACAACTCTTTGAAGGTGGAAGCAACTTTGTTTTCATATGCGCCAGACAGGAAGTG 495

```
QY 1780 GCAGAAACGACACACCATGGCGGTGAGTTTAATGGAAGGTGCAGGACTGCAAGACGGCA 1839
Db 496 GCAGACACATACTATGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCATCCAAAGGGA 555
QY 1840 TCTACGCCATGGCGCCCAAGGATGTCTTTGTGACCTGTAATATGCGCGTTACCGCGCCA 1899
Db 556 TCTATGCCATGGCCTCCCGGAGCTCTCTCTCTGAAGATCAACCTCTGCTACCGGAAGT 615
QY 1900 TGAATCTAGTCTGTCTGCGGCAAGTTCTTTGAGATTTCAGTGGCAAGGTCTTCGATCTTC 1959
Db 616 TGGGCTCTGGAAGTCTATGTGACATCTCTCGAGATCTACAAATGGGAAGCTGTTTGACCTGC 675
QY 1960 TGTCCGACAAAGCAGAAACTGCGCCTCTGGAGGATGTTAAACAGCAAGTCAGAGTGGG 2019
Db 676 TCAACAGAAAGCCAAAGCTGCGCGTCTGGAGATGGCAAGCAAGGTGCAAGTGGTGG 735
QY 2020 GACTCACCGAGAAGTGGTTCGATGGCGTGGAGAGGTACTGAAGCTCATCCAGCACGGCA 2079
Db 736 GSGTCAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATCGACATGGGCA 795
QY 2080 ATGTGCGCGCAACATCCGGCCAGAGCTGCGGCCAACTCCAAATTCGTGCGGTTCGCAAGCGG 2139
Db 796 GCGCCTGCAGAACCTCTGGGCAGACATTTGCCAACTCCAAATTCCTCCCGCTCCCAAGCGT 855
QY 2140 TTTTCCAGATTGTCTGCGCGCGCAGGCTCGACGAAGATCCATGGCAAGTTCCTGTTCA 2199
Db 856 GCTTCCAAATTAATCTTCTGAGTAAAG-----GAGAAATGATGGCAAGTTCCTTTGG 909
QY 2200 TCGATCTGGCGGCAATGAGCGGGCGTGGACACTCTCTCGGCCCATCGGCAGACGGCTA 2259
Db 910 TAGATCTGGCAGGAATGAGCGGCGGACACTTCCAGTGTCTGACCGGAGACCCGCA 969
QY 2260 TGGAGGGTCCGAGATTAACTAAATCGCTGTGTGGCCCTCAAGAGTGCATTCGTGCGTTGG 2319
Db 970 TGGAGGGCGCAAAATCAACAAGAGTCTCTTAGCCCTGAAGAGTGCATCAGGGGCCCTGG 1029
QY 2320 GCAACAGTCCGCCACTTGGCCCTCCGTGCTCCAAACTCACCAGTGTCTGGCGACT 2379
Db 1030 GACAGAACAGGCTCACACCCGTTCCGTGAGACGAGTGCATCAGAGTGTCTGAGGGACT 1089
QY 2380 GCTTCATTTGGGAGAGAGCAAGAGTGCATGATAGCCATGATCTCGCGCGGACTTAGCT 2439
Db 1090 CTTTCATTTGGGAGAACTCTAGGACTTCGATGATGTCACAGTCTCACCGAGCATAGCT 1149
QY 2440 CCTGAGACACACGCTCAACAGCTGCGCTATGCGGATCGTGTCAAGAGCTG 2492
Db 1150 CCTGTGAATATACTTTAAACACCCCTGAGATATGCAGACAGGGTCAAGGAGCTG 1202
```

## RESULT 9

```
US-10-651-510-7
; Sequence 7, Application US/10651510
; Publication No. US20040096949A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE REFERENCE: 1042A
; CURRENT APPLICATION NUMBER: US/10/651,510
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/724,215
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 09/314,464
; PRIOR FILING DATE: 1999-05-18
; PRIOR APPLICATION NUMBER: 09/594,669
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
```

```
; LENGTH: 1421
; TYPE: DNA
; ORGANISM: Human
US-10-651-510-7
```

```
Query Match 14.8%; Score 454.6; DB 7; Length 1421;
Best Local Similarity 62.6%; Pred. No. 8.3e-125;
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;
```

```
QY 1300 TGAAGAGGTGGAGCGACTGAAGGAGAATCGCGAAGACGACGCGCCCGACAGGCCGAGA 1359
Db 22 TGAAGGAAGTGGAAAAAATGAAGAAACAAGCGAGAGAAGAGGCCCGACAACTCTGAAA 81
QY 1360 TGAAGGAGGAGAAGGTGGCGCTGATGAACACAGGATCCGGGCAATCCAAACTGGAGACGG 1419
Db 82 TGAGAATGAAAGAGAG-----CTCAGGAGTATGACAGTAGTGTTCCTTCCAAACTGGGAAATTTG 135
QY 1420 CGCAAAATGATACGCGCAATATCAGAGACGCTGGGAATTTGTGCGCTGCTCGATGGCCAGG 1479
Db 136 CCCGAAATGATTAAGAAATTTGGGCTACTTTGGAAATGTCATCCACTTACTATGACTGATC 195
QY 1480 CCGTTCGATGACCAATCAGATCACAGTGTGCGTGCACAAGCGTCCCATTAGCGCGAAGGAGG 1539
Db 196 CTATCGAAGAGACACAGAAATATGTGTGTGTTAGGAAACGCCCACTGAATAAGCAAGAT 255
QY 1540 TCAATCCCAAGGAGATCGATGTCATTTTCGGTGGCGCGCAAGGACATGCTCATGTGACAG 1599
Db 256 TGGCCAAAGAAAGAAATTTGATTTCCATTCCTAGCAAGTGTCTCTCTTGGTACATG 315
QY 1600 AGCGCGCAGCAAGGTTCGACTCACAAGTTTCCTGGAGAACCAACAAGTTTCGCTTCGACT 1659
Db 316 AACCAAGTTGAAAGTGGACTTAAACAAGTATCTGGAGAAACCAAGCAATTCGCTTTGACT 375
QY 1660 ACGCTTTCAACGACACAGTGGCAAAATGCCATGGTGTATACAATATACAGACCAAGCCGTTGG 1719
Db 376 TTGCAATTTGATGAAAACAGCTTCGAATCGAAGTTGTCTACAGTTTCACAGCAAGGCCACTGG 435
QY 1720 TGAACAACATTTTCGAGGGCGGAATGGCGAGTGTCTTCGCCCTAGCGCGCAGACGGGATCGG 1779
Db 436 TACAGACAACTCTTTGAAGGTGGAAAAGCAACTTTGTTTGTGATATATGGCCAGACAGGAAGTG 495
QY 1780 GCAAAACGCAACACCATGGCGGTGAGTTTAATGAAAGGTGCAGGACTGCAAGAAACGGCA 1839
Db 496 GCAAGACACATACTATGCGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCATCCAAAGGGA 555
QY 1840 TCTACGCCATGGCGGCCAAGGATGTCTTTGTGACCCCTGAAATATGCCCGGTTTACCGGCCA 1899
Db 556 TCTATGCCATGGCGCTCCCGGAGCGTCTTCTCTCTGMAAGAAATCAACCTGCTACCGGAAGT 615
QY 1900 TGAATCTAGTCTGCTCGGCCAGTTTCTTTGAGATTTTACAGTGGCAAGGTCTTCGATCTTC 1959
Db 616 TGGGCTCTGGAAGTCTATGTGACATTTCTTCAGATCTTACAATGGGAAGCTGTTTGACCTGC 675
QY 1960 TGTCCGACAAAGCAGAAACTGCGGTCTCTGGAGATGGTAAACAGCAAGTGCAGTGTGTGG 2019
Db 676 TCAACAGAAAGGCCAAGCTGCGGTGCTGGAGATGGCAAGCAACAGTGCAGAGTGTGG 735
QY 2020 GACTCACCGAGAAGGTGGTTCGATGGCGTTCGAGGAGGTACTGAAGCTCATCCAGCACGGCA 2079
Db 736 GGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATCGACATGGGCA 795
QY 2080 ATGCTGCCCAACATCCGGCCAGACGTCGGGCCAACTCCAAATTCGTGCGGTTCGACACCGG 2139
Db 796 GCGCCTGCAGAACCTCTGGGCAGACATTTTGCCAACTCCAAATTCCTCCCGCTCCCAAGCGT 855
QY 2140 TTTTCCAGATTGTCTGCGCGCGCAGGCTCGACGAAGATCCATGGCAAGTTCCTGTTCA 2199
Db 856 GCTTCCAAATTAATCTTCGAGCTTAAGG-----GAGAAATGCAATGGCAAGTTCCTTTGG 909
QY 2200 TCGATCTGGCGGCAATGAGCGGGCGTGGACACTTCTCTCGGCCGATTCGCGAGACGGTA 2259
Db 910 TAGATCTGGCAGGAATGAGCGGCGGACACTTCCAGTGTCTGACCGGAGACCCGCA 969
```

QY 2260 TGGAGGGTCCGAGATTAAACAATCGTCTGGCCCTCAGGAGTGCATTCGTGCGTTGG 2319  
| | | | |  
Db 970 TGGAGGGCGCAGAAATCAACAAGAGTCTCTTAGCCCTCGAAGAGTGCATCAGGGCCCTGG 1029  
| | | | |  
QY 2320 GCAAAACAGTCCGCCCACTTCCCTTCCGCTCTCTCCAACTCACCCAGGTGCTCGCGACT 2379  
| | | | |  
Db 1030 GACAGAACAGGCTCACACCCGTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGGACT 1089  
| | | | |  
QY 2380 CGTTCAATTGGCAGAGAGAGCAAGACGTGCATGATAGCCATGATCTCCCGGGGACTTAGCT 2439  
| | | | |  
Db 1090 CCTTCAATTGGGAGAACTCTAGGACTTGCATGATTGCCACGATCTCACAGGCGATAAGCT 1149  
| | | | |  
QY 2440 CCTGCGAGCACGCTCAACACCGCTGGCTATCGGATCGTGTCAAGGAGCTG 2492  
| | | | |  
Db 1150 CCTGTGAATATACTTTAAACACCCCTGAGATATGCACAGGGGTCAAGGAGCTG 1202  
| | | | |  
RESULT 10  
US-10-651-510-13  
; Sequence 13, Application US/10651510  
; Publication No. US20040096949A1  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: Novel motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1042A  
; CURRENT APPLICATION NUMBER: US/10/651,510  
; CURRENT FILING DATE: 2003-08-28  
; PRIOR APPLICATION NUMBER: US/09/724,215  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/295,612  
; PRIOR FILING DATE: 1999-04-20  
; PRIOR APPLICATION NUMBER: 09/314,464  
; PRIOR FILING DATE: 1999-05-18  
; PRIOR APPLICATION NUMBER: 09/594,669  
; PRIOR FILING DATE: 2000-06-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 2172  
; TYPE: DNA  
; ORGANISM: Human  
US-10-651-510-13  
Query Match 14.8%; Score 454.6; DB 7; Length 2172;  
Best Local Similarity 62.6%; Pred. No. 1.1e-124;  
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;  
QY 1300 TGAAGAGGTGGAGCGACTGAAGGAGAAATCGCGAGAGCGACGCGCCGACAGCGCGAGA 1359  
| | | | |  
Db 578 TGAAGGAAAGTGGAAAAAATGAAGAAACAAGCGAGAGAGAAAGAGGCCAGAACTCTGAAA 637  
| | | | |  
QY 1360 TGAAGGAGGAGAGGTGGCGCTGATCAACACGGATCCGGGCAATCCAACTCGGAGACGG 1419  
| | | | |  
Db 638 TGAAGATGAAGAGAG-----CTCAGAGATGACAGATGATTTTCCAAACTGGGAATTG 691  
| | | | |  
QY 1420 CGCAATGATACCGCAATATCAGAGCACGCTGGAATTTGTGCGCTGCTCGATGGCCAGG 1479  
| | | | |  
Db 692 CCGCAATGATTAAGAAATTTGGGCTACTTTGGAATGTCATCCACTTACTATGACTGATC 751  
| | | | |  
QY 1480 CGGTGATGACCATCAGATCAGAGTGTGCGCGCAAGCGTCCCATTAGCCGCAAGGAGG 1539  
| | | | |  
Db 752 CTATCGAAGAGCACAGATATGTCTGTGTAGGAAAGCCCACTGTAATAAGCAAGAAAT 811  
| | | | |  
QY 1540 TCAATCGCAGGAGATCGATGTCATTTCCGTCCGCGGAGGACATGCTCATCGTGCAG 1599  
| | | | |  
Db 812 TGGCCAAAGAAAGAAATTTGATGATTTCCATTCTTAGCAAGTGTCTCTCTTTGGTACATG 871  
| | | | |  
QY 1600 AGCCGGCGCAGCAGGTGCGACCTCACCAAGTTCCTGGAGAACCAAGATTTGCTTCGACT 1659  
| | | | |  
Db 872 AACCCAAAGTTGAAGTGGACTTAACAAAGATATCTGGAGAACCAAGCATTTCTGCTTGACT 931  
| | | | |

QY 1660 AGCCCTTCAACGACAGCGTCCGACAAATGCGCATGGTATATAAAATACACAGCCAAAGCCCTTGG 1719  
| | | | |  
Db 932 TTGCATTTTATGAAAAAGCTTTCGAATGAAGTTTGTCTACAGGTTTACAGCAAGGCCACTGG 991  
| | | | |  
QY 1720 TGAATAACCAATTTTCAGAGGGCGGAATGCGCAGCGTCTTCGCTTACGGCCAGAGCGGATCGG 1779  
| | | | |  
Db 992 TACAGACAATCTTTGAGGTGGAAAGCAACTTTTTCATATATGCCAGACAGGAAGTG 1051  
| | | | |  
QY 1780 GCAAAAACGACACCAATGCGCGGTGAGTTTAAATGGAAGGTGCGAGGACTGCAAGAACGGCA 1839  
| | | | |  
Db 1052 GCAAGACACATACTATGTCGCGGAGACCTCTCTGGGAAAGCCCAAGATGCAATCCAAAGGA 1111  
| | | | |  
QY 1840 TCTACGCCATGCGGGCCAAAGGATGCTTTGTGACCTGTAATATGCGCGCTTACCGCGCCA 1899  
| | | | |  
Db 1112 TCTATGCCATGCGCTTCCCGGAGCGTCTTCTCTCTGGAAGAAATCAACCTCTGCTACCGAAGT 1171  
| | | | |  
QY 1900 TGAATCTAGTCTGCTCGGCGAGTCTTTCTTTGAGATTTTACAGTGGCAAGGTCTTCGATCTTC 1959  
| | | | |  
Db 1172 TGGGCTTGAAGTCTATGTGACATTTCTTCGAGATCTACATGGGAAGCTGTTTGACCTGC 1231  
| | | | |  
QY 1960 TGTCCGACAAAGCAGAAACTGCGCGTCTCTGGAGGATGGTAAACAGCAAGTGCAGGTGGTGG 2019  
| | | | |  
Db 1232 TCAACAAAGAAAGGCCAAGCTGCGCGTCTGGAGGATGGCAAGCAACAGGTGCAAGTGGTGG 1291  
| | | | |  
QY 2020 GACTCACCGAGAAAGTGGTGGATGGGTCGAGGAGTACTGAAGCTCATCCAGCAGCGCA 2079  
| | | | |  
Db 1292 GGGTCGAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGATCGACATGGCA 1351  
| | | | |  
QY 2080 ATGCTGCCGCAACATCCGSCCAGAGCTGGGCCAAGCTCCAAATTCGTCGCGTTCGACGCGG 2139  
| | | | |  
Db 1352 GCGCTGACAGAACTCTGGGACAGATTTGCCAACTCCAAATTCCTCCCGCTCCACGCGT 1411  
| | | | |  
QY 2140 TTTTCCAGATTTGCTGCGCGCGCAGGGCTCGACGAAAGATCCATGGCAAGTTCCTCGTTCA 2199  
| | | | |  
Db 1412 GCTTCCAAATTAATCTTCGAGCTAAAG-----GAGATGATGCAAGTTCCTCTTGG 1465  
| | | | |  
QY 2200 TCGATCTGCGGGCAATGAGCGGGCGTGGACACTTCTTCGGCGGATCGGCGAGACGCGTA 2259  
| | | | |  
Db 1466 TAGATCTGGCAGGAAATGAGCGGAGCGCGACACTTCCAGTGTGACCGGCGAGACCGCA 1525  
| | | | |  
QY 2260 TGGAGGGTCCGAGATTAACAATCGCTGCTGCGCCCTCAAGGAGTGCATTCGTGCGTTGG 2319  
| | | | |  
Db 1526 TGGAGGGCGCAGAAATCAACAAGATCTTTAGCCCTTGAGGAGTGCATCAGGGCCCTGG 1585  
| | | | |  
QY 2320 GCAAAACAGTCCGCCCACTTTGCCCTTCCGTCTTCCAACTCACCCAGGTGCTCGCGACT 2379  
| | | | |  
Db 1586 GACAGAACAAAGCTCACACCCGTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGGACT 1645  
| | | | |  
QY 2380 CATTCAATTGGCGAGAGAGAGAGCTGTCATGATAGCCATGATTCGCCGGGACTTAGCT 2439  
| | | | |  
Db 1646 CTTTCAATTGGGAGAACTCTAGGACTTGCATGATTTGCCACGATCTCACCGGCGATAAGCT 1705  
| | | | |  
QY 2440 CTGCGAGCACAGCTCAACACGCTGCGCTATGCGGATCGTCTCAAGGAGCTG 2492  
| | | | |  
Db 1706 CCTGTGAATATACTTTAAACACCCCTGAGATATGCAGACAGGGTCAAGGAGCTG 1758  
| | | | |  
RESULT 11  
US-10-240-965-208  
; Sequence 208, Application US/10240965  
; Publication No. US20030165924A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: SHIFFMAN, Dov  
; APPLICANT: SOMOGYI, Roland  
; APPLICANT: LAWN, Richard M.  
; APPLICANT: SEILHAMER, Jeffrey J.  
; APPLICANT: PORTER, Gordon J.  
; APPLICANT: MIKITA, Thomas  
; APPLICANT: TAI, Julie  
; TITLE OF INVENTION: GENES EXPRESSED IN FOAM CELL DIFFERENTIATION  
; FILE REFERENCE: PA-0025 PCT  
; CURRENT APPLICATION NUMBER: US/10/240,965



```
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: 60/195,106
; PRIOR FILING DATE: 2000-04-05
; NUMBER OF SEQ ID NOS: 276
; SOFTWARE: PERL Program
; SEQ ID NO 208
; LENGTH: 2865
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030165924A1 331025.1
US-10-240-965-208

Query Match
Best Local Similarity 14.7%; Score 453; DB 6; Length 2865;
Matches 746; Conservative 0; Mismatches 435; Indels 12; Gaps 2;

QY 1300 TGAAGAGGTGGAGGAGCTGAAGGAGAAATCGGAGAGCGAGCGCGCCGACAGAGCGGAGG 1359
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
675 TGAAGGAGTGGAAAAATGAAGAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 734
QY 1360 TGAAGGAGAGAGAGTGGCGCTGATGAACAGAGATCCGGGGCAATCAAACTGGGAGAGCG 1419
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
735 TGAGAAATGAAGAGAG-----CTCAGGAGATATGACAGTAGTATTTTCCAAACTGGGAATTTG 788
QY 1420 CCAAAATGATAGCGCAATATCAGACGACGCTGGAAATTTGTCCGCTGCTCGATGCCAGG 1479
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
789 CCGGAATGATTAAGAAATTCGGGTACTTTGGAATGTCATCCACTTACTATGACTGATC 848
QY 1480 CCGTCGATGACCATCAGATCAGTGTGCGCAAGCGTCCCAATAGCCGCAAGGAGG 1539
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
849 CTATCGAAGAGCACAGATATGTGTCTGTTAGGAACGCCCACTGATTAAGCAAGAT 908
QY 1540 TCAATCGAAGAGATCGATGCTATTTTCGTCGCCCGCAAGACATGCTCATCGTGACG 1599
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
909 TGGCAAGAAAGAAATGATGATGATTTCCATTCCTAGCAAGTGTCTCTCTTGGTACATG 968
QY 1600 AGCCGCGCAGCAAGTTCGACCTCACCAAGTTCTCGGAGAACACAAAGTTTCGCTCGACT 1659
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
969 AACCAAGTTGAAAGTGAATTAACAAAGATCTGGAGAACCAAGCATCTGCTTTGACT 1028
QY 1660 AGCCTTCAACGACACGTGCGCAATGCCATGGTATACAAATACACAGCCCAAGCGTTGG 1719
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1029 TTGCATTTGATGAACACAGCTTCGAATGAAGTTGCTACAGGTTACACAGAGCCACTGG 1088
QY 1720 TGAAGAACATTTTCAGGCGGAGATGGCGAGCTGCTTCGCTACGCGCAGACGGGATCGG 1779
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1089 TACAGACAATCTTTGAAGGTGGAAAGCAACTTGTGTCATATGCGCCAGACAGGAAGTG 1148
QY 1780 GCMAAACGCACACCATGGCGGTGATTTAATGGAAGGTGCAGGACTGCAGAACGCA 1839
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1149 GCAAGACACATCTATGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCATCCAAAGGGA 1208
QY 1840 TCTAGCCATGGCGCCCAAGGATGCTTTGTGACCCCTGAATATGCGCGTTTACCGCGCA 1899
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1209 TCTATGCCATGGCTCCCGGAGCTCTCTCTCTGAAGAAATCAACCTGCTACCGGAGT 1268
QY 1900 TGAATCTAGTCTGCTCGGCCAGTTTCTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTC 1959
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1259 TGGGCTTGAAGTCTATGTGACATCTCTCGAGATCTACAATGGGAAGCTGTTTGCACCTGC 1328
QY 1960 TGTCCGCAACGACAGAAATCGCGCTCTGGAGGATGTTAAACAGCAAGTGCAGGTGGTGG 2019
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1329 TCAACAGAGAGGACAGCTGCGCTGCTGGAGGAGCGGCAACAGGTCAAGTGGTGG 1388
QY 2020 GACTCACCGAAGGTGGTTCGATGCGCTCGAGGAGGTACTGAAGTCTATCCAGCACGCGCA 2079
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1389 GGCTGCAGGAGCATCTGCTTAACTCTGCTGATGATGTCATCAAGATGATGCATGGGCA 1448
QY 2080 ATGTGCGCCGAACATCCGCGCAGAGCTCGGCCAACTCCAAATTCGCGGTTTCGACGCGG 2139
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1449 GCGCTTGAGAAACCTCTGGGAGAGCATTTGCGCAACTCCAAATTCCTCCGCTCCACACGCT 1508
```

```
QY 2140 TTTTCCAGATTGTCTCGCGCGCAGGCTCGAGGAAGATCCATGCAAGTTCTCGTTCA 2199
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1509 GCTTCCAAATTAATCTTCGAGCTAAAGG-----GAGAAATGCAAGTTCTCTTTGG 1562
QY 2200 TCGATCTGGCGGCAATGAGCGGGCGGTGGACACTTCTCGGCCGATCGGAGACGCGTA 2259
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1563 TAGATCTGGCAGGGAATGAGCGGCGGACACTTCCAGTGTGACCGGAGACCGCA 1622
QY 2260 TGGAGGGTGGCGAGATTAAACAAATCGTGTGGCCCTCAAGAGTGCATTCGTGCTGG 2319
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1623 TGGAGGGCGCAGAAATCAACAAAGAGTCTCTTAGCCCTGAAGAGTGCATCAGGCGCCTGG 1682
QY 2320 GCAAAACAGTGGCCCACTTGGCCCTCGTGTCTCCAAACTCACCAGGTGCTGCGGACT 2379
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1683 GACAGAAACAGGCTCACACCCCTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGACT 1742
QY 2380 CGTTCAATGGCGAGAGCAAGACGTGCATGATAGCCATGATCTCGCGGGACTTAGCT 2439
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1743 CTTCAATGGGAGAACTTAGGACTTGCATGATGCGACGATCTCACAGGCAATAGCT 1802
QY 2440 CTTGCGAGCACACCTCAACACGCTGCGCTATGCGGATCGGTCAAGGAGCTG 2492
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1803 CCTGTGAATATATCTTTAAACACCCCTGAGATATGACAGAGGTCAAGGAGCTG 1855
```

```
RESULT 12
US-10-510-3
; Sequence 3, Application US/10651510
; Publication No. US20040096949A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Novel motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1042A
; CURRENT APPLICATION NUMBER: US/10/651,510
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/724,215
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 09/314,464
; PRIOR FILING DATE: 1999-05-18
; PRIOR APPLICATION NUMBER: 09/594,669
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1175
; TYPE: DNA
; ORGANISM: Human
US-10-651-510-3
```

```
Query Match 14.6%; Score 449.8; DB 7; Length 1175;
Best Local Similarity 64.0%; Pred. No. 2.1e-123;
Matches 698; Conservative 0; Mismatches 387; Indels 6; Gaps 1;

QY 1402 ATCAAACTGGGAGACGGCGCAATGATACGCAATATCAGAGCACGCTGGAATTTGTGC 1461
Db      1 ATGCAAACTGGGAATTTGCGCGAATGATTAAAGAAATTCGGGCTACTTTGGAATGTCTC 60
QY 1462 CGGTGCTCGATGGCGAGCGCTCGATGACCATCAGATCAGATGTCGTCGCGGAGGCTC 1521
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 CACTTACTATGATGATCTCTATCGAAGAGCACAGAAATATGTGCTGTGTGTAGGAAACGCC 120
QY 1522 CCATTAGCCGCAAGGAGGTCAATCGCAAGGAGATCGATGTCATTTCCGTGCGCGCAAGG 1581
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121 CACTGATAGCAAGAAATTTGGCCCAAGAAATGATGATTTCCATTTCTTAGCAAGT 180
QY 1582 ACATGCTCATCTGTCAGAGCGCGGAGCAAGGTTCAGACCTCAACAAAGTTCTGAGAAC 1641
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181 GTCTCTCTTGGTACATGAACCAAGTTGAAAGTGGACTTAAACAAAGTATCTGAGAAC 240
```

```
QY 1642 ACAAGTTTCGCTCGACTACGCTTCAACGACACGTCGCAATGCGCATGGTATACAAAT 1701
Db 241 AAGCATTTCTGCTTGAATTTGCAATTTGATGAACACAGCTTCCGAATGAAGTTGTCTACAGGT 300
QY 1702 ACACAGCCAAAGCCGTTGGTGAACAAACATTTTCAGAGGCGGAATGGCGACGTCTTCGCCCT 1761
Db 301 TCACAGCAAGCCACTGCTACAGCAATCTTTGAGGTGGAAGCAATCTGTTTGTGCAT 360
QY 1762 ACGCCAGACGGGATCGGCGAACAACGACACCATATGGCGCGTGTGATTTAATGGAAGGTGC 1821
Db 361 ATGGCCAGACAGGAAGTGGCAAGACACATATATGGCGGAGACCTCTCTGGGAAGCCC 420
QY 1822 AGCACTGCAAGACGGCATCTAGCCCATGGCGCCGAAGATGTCTTTGTGACCTGAATA 1881
Db 421 AGAATGCATCAAGAGGATCTATGCCATGGCTCCCGGAGCGTCTTCTCTGAAGAATC 480
QY 1882 TGGCGGCTTACCGCGCCATGAATCTAGTCTGCTCGGCCAGTTCTTTGAGATTTACAGTG 1941
Db 481 AACCTGTCTACCGAAGTTGGGCTGTGAAGTCTATGTGACATTTCTCGAGATCTCAATG 540
QY 1942 GCAAGTCTTTCGATCTTCTGTCGCAAGCAGAAACTGCGCGTCTCTGAGGATGGTAAAC 2001
Db 541 GGAAGCTGTTGACCTGCTCAACAAGAGCCAAAGCTGCGCGTCTGGAGATGGCAAGC 600
QY 2002 AGCAAGTGCAGGTGGGACTCACCGAAGGTGTGATGGCGGTGCGAGGATGACTGA 2061
Db 601 AACAGTGAAGTGTGGGCTGTCAGAGCATCTGGTTAACTCTGCTGATGATGTATCA 660
QY 2062 AGCTCATCCAGCAGCAATGCTGCCGCAACATCCCGCCAGAGCTGCGCGTCTGCAAT 2121
Db 661 AGATGATCAGATGGGCGGCTGCGAGAACCTCTGCGGAGACATTTGGCAACTCCAAAT 720
QY 2122 CGTCCGCTTCGACGCGCTTTTCCAGATTTGCTGCGGCGCGCAGGCTCGACGAAGTCC 2181
Db 721 CCTCCGCTCCACGCGTCTTCCAAATTTATCTTCGAGCTAAAG-----GAGATGC 774
QY 2182 ATGGCAAGTTCTGTTTCATGATCTGGCGGCAATGAGCGGCGGTGGAACATCTCTCGG 2241
Db 775 ATGGCAAGTTCTTCTTGGTAGATCTGGCAGGGAATGAGCGGCGGCACTTCCAGTG 834
QY 2242 CGATCGGCGAGCGCTATGAGGCGCGAGATTAACAAATCGCTGCTGGCCCTCAAG 2301
Db 835 CTGACCGGAGACCCGATGAGGCGCGAGAAATCAACAGAGTCTCTTAGCCCTGAAG 894
QY 2302 AGTGCAATTCGTGGTGGGCAACAGTCGCGCCACATTTGCCCTTCGTGTCTCAAACTCA 2361
Db 895 AGTGCAATCAGGCGCTGGGACAGAACAGGCTCACACCCGTTCCGTGAGAGCAAGCTGA 954
QY 2362 CCCAGTGTGCGGACTCGTTTCATTTGGGAGAGCAAGACGTCGATGATAGCCATGA 2421
Db 955 CACAGTGTGAGGACTCTTTCATTTGGGAGAACTCTAGGACTTGCATGATTTGCCACGA 1014
QY 2422 TCTCGCGGACTTACTCTCGGACACACGCTCAACAGCTCGGCTATGCGATGCGTGTG 2481
Db 1015 TCTCACAGGCAAGCTCTCTGTGAATATCTTTAAACCCCTGAGATATGACAGAGG 1074
QY 2482 TCAAGGAGCTG 2492
Db 1075 TCAAGGAGCTG 1085
```

## RESULT 13

```
US-10-651-510-9
; Sequence 9, Application US/10651510
; Publication No. US20040096949A1
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: Novel motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1042A
; CURRENT APPLICATION NUMBER: US/10/651,510
```

```
; CURRENT FILING DATE: 2003-08-28
; PRIOR APPLICATION NUMBER: US/09/724,215
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 09/314,464
; PRIOR FILING DATE: 1999-05-18
; PRIOR APPLICATION NUMBER: 09/594,669
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1304
; TYPE: DNA
; ORGANISM: Human
; US-10-651-510-9
```

```
Query Match 14.5%; Score 449.8; DB 7; Length 1304;
Best Local Similarity 64.0%; Pred. No. 2.2e-123;
Matches 698; Conservative 0; Mismatches 387; Indels 6; Gaps 1;
```

```
QY 1402 ATCCAACTGGGAGACGGCGCAAAATGATACGCGAATATCAGAGCAGCTGGAAATTTGTGC 1461
Db 1 ATGCAAACTGGGAATTTGCCGAATGATTAAGAATTTTCGGCTACTTTTGGAAATGTCATC 60
QY 1462 CGCTGCTCGATGGCCAGGCGCTCGATGACCATCAGATCAGATGTCGTGCGCAAGCGTC 1521
Db 61 CACTTACTATGACTGATCTCTATCGAAGACACAGATATGTCGTCTGTAGGAACGCC 120
QY 1522 CCATTAGCCGCAAGAGGTCAATCGGAGAGATCATGTCAATTCGGTGCCTCGCGCAAGG 1581
Db 121 CACTGAATAAGCAAGAAATTTGGCCAAAGAAATTTGATGTGATTTCCATTTCTTAGCAAGT 180
QY 1582 ACATGCTCATCTGTGACGAGCGCGCAGCAAGCTCGACCTCACCAAGTTCTCGGAGAAC 1641
Db 181 GTCTCTCTTGTGTACATGAACCAAGTTGAAAGTGAGCTTAACAAAGTATCTCGAGAAC 240
QY 1642 ACAAGTTTCGCTTCGACTACGCTTCAACGACACGTCGCACAATGCGCATGGTATACAAAT 1701
Db 241 AAGCATTTCTGCTTTCGCTTTGATGATGAACAGCTTCGAATGAAGTTGTCTACAGGT 300
QY 1702 ACACAGCCAAAGCGTTGGTGAACCAATTTTCAGGCGCGAATGGCGACGTCTTCGCCCT 1761
Db 301 TCACAGCAGGCGCTCTGGTACAGCAATCTTTGAAAGTGGAAAGCAACTGTGTTTGTGCAT 360
QY 1762 ACGGCCAGACGGGATCGGSCAAACGACACCATGCGGCGTCAAGTTAATGGAAGGTGC 1821
Db 361 ATGGCCAGACAGGAAGTGGCAAGACACATCTATGCGGCGAGACCTCTCTCGGAAAGCCC 420
QY 1822 AGGACTGCAAGAACGCGCATCTACGCCATGGCGGCCAAGGATGCTTTTGTGACCTCAATA 1881
Db 421 AGAATGCATCCAAAGGATCTATGCCATGGCTCCCGGACGCTTCTCTCTGAAGATC 480
QY 1882 TGCCCGGTTACCGCGCCATGAATCTAGTGTCTCGGCGATTTCTTTTGAAGTTACAGTG 1941
Db 481 AACCCCTGTCTACCGGAAGTTGGGCTTGAAGTCTATGTGACATTTCTTCGAGATCTACAATG 540
QY 1942 GCAAGTCTTCGATCTTCTGTCGCAAGCAGCAAGACTCGCGCTCTCGGAGGATGGTAAC 2001
Db 541 GGAAGCTGTTTGAATCTGCTCAACAAAGAGGCCAAGCTGCGCTGTGAGGATGGCAAGC 600
QY 2002 AGCAAGTGCAGGTGGGACTCACCGAAGGTTGTCATGGCGCTCGAGGAGGTACTGA 2061
Db 601 AACAGTGCAGGTGGTGGGCTGTCAGGAGCATCTGGTTAACTCTGCTGATGATGTATCA 660
QY 2062 AGCTCATCAGCAAGGCAATGTCGCCGCAACATCCCGCCAGAGCTGCGCGCAACTCCAAAT 2121
Db 661 AGATGATCAGATGGGCGGCTGCGAGAACCTCTGCGGAGACATTTTGCACACTCCAAAT 720
QY 2122 CGTCCGCTTCGACGCGCTTTTCCAGATTTGTCGCGCGCGGAGGCTCGAGAGATCC 2181
Db 721 CCTCCGCTCCACGCGTCTTCCAAATTTATCTTCGAGCTAAAG-----GAGATGC 774
```



Db 1766 CCTGTGATATACCTTTAAACACCCCTGAGATATGACAGACAGGGTCAAGGAGCTG 1818

## RESULT 15

US-09-954-456-710  
; Sequence 710, Application US/09954456  
; Patent No. US20020115057A1  
; GENERAL INFORMATION:  
; APPLICANT: Young, Paul  
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand  
; FILE REFERENCE: 689290-76  
; CURRENT APPLICATION NUMBER: US/09/954,456  
; CURRENT FILING DATE: 2001-09-18  
; PRIOR APPLICATION NUMBER: US/60/233,617  
; PRIOR FILING DATE: 2000-09-18  
; PRIOR APPLICATION NUMBER: US/60/234,052  
; PRIOR FILING DATE: 2000-09-20  
; PRIOR APPLICATION NUMBER: US/60/234,923  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,134  
; PRIOR FILING DATE: 2000-09-25  
; PRIOR APPLICATION NUMBER: US/60/235,637  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,638  
; PRIOR FILING DATE: 2000-09-26  
; PRIOR APPLICATION NUMBER: US/60/235,711  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,720  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,840  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: US/60/235,863  
; PRIOR FILING DATE: 2000-09-27  
; NUMBER OF SEQ ID NOS: 2276  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 710  
; LENGTH: 2740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-954-456-710

Query Match 14.6%; Score 449.8; DB 3; Length 2740;  
Best Local Similarity 62.4%; Pred. No. 3.4e-123;  
Matches 744; Conservative 0; Mismatches 437; Indels 12; Gaps 2;  
QY 1300 TGAACAGGTGGAGGACTGAAGGAGAAATCGGAGAGCGACGCGCCGACAGCGGCGAGA 1359  
Db 638 TGAGGAGTGGAAAAATGAAGAACAGCGAGAGAGAGAGAGGCGCCAGAACTCTGA 697  
QY 1360 TGAAGGAGGAGAGGTGGCGCTGATGAACACAGGATCCGGGCAATCCAAACTCGGAGACGG 1419  
Db 698 TGAGATGAAGAGAG-----CTCAGGATATGACAGTAGTTTCCAACTGGGAATTG 751  
QY 1420 CGCAATGATACCGGAATATCAGAGACGCTGGAATTTGTGCGGTGCTCGATGGCCAGG 1479  
Db 752 CCGGAATGATTAAGAAATTCGGGCTACTTTGGAATGTCTATCCACTTACTATGACTGATC 811  
QY 1480 CCGTCGATCAGATCAGATGCTGCGTGGCAGAGGCTCCCATAGCCGCAAGGAGG 1539  
Db 812 CTATCGAAGAGCAGAGATATGTGTCTGTGTAGGAAACGCCCACTGAATTAAGCAAGAT 871  
QY 1540 TCAATCGAAGGAGATCGATGTCTATTTCCGTGCGCGCAAGGACATGCTCATCGTGACG 1599  
Db 872 TGGCCAAAGAAAGAAATTTGATGTATTTCCATTCCTAGCAAGTGTCTCTCTTGGTACATG 931  
QY 1600 AGCCGCGCAGCAGGTGCGACCTCAACAGTTCTGTGAGAACCAAGTTTCTGCTTGACT 1659  
Db 932 AACCCTTCAACGACGCTGCGACATGCGATGGTATACAAATACACAGCCCAAGCCGTTGG 1719  
QY 1660 AGCCCTTCAACGACGCTGCGACATGCGATGGTATACAAATACACAGCCCAAGCCGTTGG 1719  
Db 992 TTGCAATTTGATGAACAGCTTCGAATGAAGTTGTCTACAGGTTTACAGGTTTCAAGGCACTGG 1051

QY 1720 TGAACACCACTTTTCGAGGCGGGAATGGCGACGCTGCTTTCGCTACGCGCCAGACGGGATCGG 1779  
Db 1052 TACAGACAATCTTTGAAGGTGGAAGCAACTTGTGTTTGCATATGCGCAGACAGGAGTG 1111  
QY 1780 GCAAAACGACACCATGGCGGTGAGTTTAAATGGAAGGTGAGAGCTGCAAGAACGGCA 1839  
Db 1112 GCAAGACACATATGAGGCGGAGACCTCTCTGGAAGCCAGAAATGCAATCCAAAGGGA 1171  
QY 1840 TCTACGCCATGGCGGCAAGGATGCTTTGTGACCTGACCTGAATATGCGCGTTACCGCGCCA 1899  
Db 1172 TCTATGCCATGGCTCCCGGAGCTTCTCTCTGAAGAAATCAACCTGCTACCGAAGT 1231  
QY 1900 TGAATCTAGTCTGCTCGGCAGTTTCTTTGAGATTTTACAGTGGCAAGGTCTTCGATCTTC 1959  
Db 1232 TGGGCTCGGAAGTCTATGTGACATTTCTCGAGATCTACAATGGGAAGCTGTTTGAACCTG 1291  
QY 1960 TGTCCGACAGCAGAACTGCGGCTCTGAGGATGTTAAACAGCAAGTGCAGGTGGTGG 2019  
Db 1292 TCAACAAAGAAAGCCAAAGCTGCGGCTGCTGAGGACGGCAAGCAAGGTGCAAGTGGTGG 1351  
QY 2020 GACTCACCGAGAAAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2079  
Db 1352 GGCTGCAGAGGATCTGGTTAACTCTGCTGATGATGTCATCAAGATGCTCGACATGGGCA 1411  
QY 2080 ATGCTGCCGCAACATCCGGCCAGAGCTCGGCCCAACTCCAAATTCGTCGCGTTCCAGCCG 2139  
Db 1412 GCGCTCGCAAGACCTCTCGGCGAGACATTTTGGCAACTCCAAATTCCTCCCGCTCCACGCGT 1471  
QY 2140 TTTTCCAGATTTGCTGCGGCGCGAGGCTCGACGAGATCCATGGCAAGTTCTCTGTTCA 2199  
Db 1472 GCTTCCAAATTTATTTCTCGAGCTAAAGG-----GAGAAATGCAATGGCAAGTTCTCTTTGG 1525  
QY 2200 TCGATCTGCGCGGCAATGAGCGGGGCTGGACACTTTCCTCGCGCGATCGGACAGCGGTA 2259  
Db 1526 TAGATCTGCGAGGAAATGAGCGGCGAGACACTTTCAGTGTCTGACCGGAGCCCGCA 1585  
QY 2260 TGGAGGTGCGGAGATTAACAAAATCGCTGCTGCGCCCTCAAGGAGTGCATTTCGTGCGTTGG 2319  
Db 1586 TGGAGGCGCAGAAATCAACAAGAGTCTCTTAGCCCTGAAGGAGTGCATCAGGGCCCTGG 1645  
QY 2320 GCAACAGTGGCGCCCACTTGGCCCTTCCGCTGCTCCAACTCACCGAGGTGCTGCGGACT 2379  
Db 1646 GACAGAACAAAGGCTCACACCCGTTCCGTTGAGAGCAAGCTGACACAGGTTGCTGAGGACT 1705  
QY 2380 CGTTCAATGGCGAGAGAGCAAGAGCTGCATGATAGCCATGATCTGCGCGGACTTAGCT 2439  
Db 1706 CTTTCAATGGGAGAACTCTAGGACTTGCATGATTTGCCACGATCTCACCAGGCAATAGCT 1765  
QY 2440 CTTGCGAGCACACGCTCAACACGCTGCGCTATGCGGATCGTGTCAAGGAGCTG 2492  
Db 1766 CTTGTGAATATACTTTAAACACCCCTGAGATATGACAGACAGGCTCAAGGAGCTG 1818

Search completed: May 12, 2006, 22:17:06  
Job time : 2384 secs

***This Page Blank (uspto)***

15	52.4	1.7	3330	11	US-10-9332-182A-84	Sequence 84, Appl
16	52.4	1.7	3330	11	US-10-9332-182A-84	Sequence 84, Appl
17	51.6	1.7	3304	13	US-10-979-095-19	Sequence 19, Appl
18	48.4	1.6	3336	11	US-10-9332-182A-75421	Sequence 75421, A
19	48.6	1.6	3336	11	US-10-9332-182A-75421	Sequence 75421, A
20	47.8	1.6	1515	7	US-09-925-068A-5627	Sequence 5627, Ap
21	47.8	1.6	1515	11	US-10-301-480-106864	Sequence 106864, Ap
22	47.8	1.6	1515	12	US-10-301-480-720273	Sequence 720273, Ap
23	45.8	1.5	1703	10	US-10-995-561-160	Sequence 160, App
24	45.8	1.5	1949	10	US-10-995-561-161	Sequence 161, App
25	45.8	1.5	1956	10	US-10-995-561-157	Sequence 157, App
26	45.8	1.5	2781	10	US-10-995-561-285	Sequence 285, App
27	45.8	1.5	25968	10	US-10-995-561-13248	Sequence 13248, A
28	45.8	1.5	28933	10	US-10-995-561-13285	Sequence 13285, A
29	45.2	1.5	2871	11	US-10-330-773-708	Sequence 708, App
30	44.2	1.4	568	7	US-09-925-068A-119416	Sequence 119416, Ap
31	44.2	1.4	590	7	US-09-925-068A-508385	Sequence 508385, Ap
32	43.8	1.4	1095	11	US-10-9332-182A-82043	Sequence 82043, A
33	43.8	1.4	1095	11	US-10-9332-182A-82043	Sequence 82043, A
34	43.6	1.4	1331	17	US-11-136-527-2116	Sequence 2116, Ap
35	43.6	1.4	1474	18	US-11-095-568A-16441	Sequence 16441, A
36	43.6	1.4	2261	18	US-11-051-720-72	Sequence 72, Appl
37	43.6	1.4	2945	18	US-11-051-720-71	Sequence 71, Appl
38	43.6	1.4	6433	11	US-10-330-773-711	Sequence 711, App
39	43.6	1.4	6927	13	US-10-960-414-231	Sequence 231, App
40	43.6	1.4	7004	18	US-11-051-720-66	Sequence 66, Appl
41	43.6	1.4	7128	18	US-11-051-720-65	Sequence 65, Appl
42	43.6	1.4	7855	11	US-10-330-773-501	Sequence 501, App
43	43.4	1.4	2077	17	US-11-136-527-548	Sequence 548, App
44	43.4	1.4	2115	11	US-10-9332-182A-3350	Sequence 3350, Ap
45	43.4	1.4	2115	11	US-10-9332-182A-3350	Sequence 3350, Ap

[illegible]

## ALIGNMENTS

RESULT 1  
US-11-128-061-3433  
; Sequence 3433, Application US/11128061  
; Publication No. US20060003958A1

; APPLICANT: Melville, Mark W.  
 ; APPLICANT: Charlebois, Timothy S.  
 ; APPLICANT: Mounts, William M.  
 ; APPLICANT: Hann, Louane E.  
 ; APPLICANT: Sinacore, Martin S.  
 ; APPLICANT: Leonard, Mark W.  
 ; APPLICANT: Brown, Eugene L.  
 ; APPLICANT: Miller, Christopher P.

```

, TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
, TITLE OF INVENTION: TO MONITOR GENE EXPRESSION
, FILE REFERENCE: 01997.027701
, CURRENT APPLICATION NUMBER: US/11/138,061
, CURRENT FILING DATE: 2005-05-11
, PRIOR APPLICATION NUMBER: US 60/570,425
, PRIOR FILING DATE: 2004-05-11
, NUMBER OF SEQ ID NOS: 7285
, SOFTWARE: Patent In version 3.3

```

```

; SEQ ID NO 3433
; LENGTH: 2703
; TYPE: DNA
; ORGANISM: Cricetulus griseus
US-11-128-061-3433

```

Query Match 15.0%; Score 461.4; DB 17; Length 2703;  
Best Local Similarity 61.5%; Pred. No. 1.le-107;  
Matches 779; Conservative 0; Mismatches 476; Indels 12;

Qy	1226	541	Db	600
	GCGGGCACAGGAGTGC	GCGGGCACAGGAGTGC	GCGGGCACAGGAGTGC	GCGGGCACAGGAGTGC
	CGCCATGCGGGCCAC	CGCCATGCGGGCCAC	CGCCATGCGGGCCAC	CGCCATGCGGGCCAC
	CAACCCAGGAGCTGG	CAACCCAGGAGCTGG	CAACCCAGGAGCTGG	CAACCCAGGAGCTGG
	CGCGCGCTAGTACCG	CGCGCGCTAGTACCG	CGCGCGCTAGTACCG	CGCGCGCTAGTACCG

QY 1286 GCGATCGCACGCATTGAAAGAGGTGGAGCGGACTGAAGGAGAAATCGCGAGAACGACGCGC 1345

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	†				
1	461.4	15.0	2703	17	US-11-128-061-3433		Sequence 3433, Ap
2	461.4	15.0	2703	17	US-11-128-049-3433		Sequence 3433, Ap
3	453	14.7	2832	14	US-11-145-307A-22		Sequence 22, Appl
4	449.8	14.6	2740	10	US-10-955-054A-137		Sequence 137, App
5	449.8	14.6	2740	10	US-10-955-054A-101		Sequence 101, App
6	449.8	14.6	2740	13	US-10-960-A14-299		Sequence 299, App
7	243	7.9	1400	17	US-11-128-061-7075		Sequence 7075, Ap
8	243	7.9	1400	17	US-11-128-049-7075		Sequence 7075, Ap
9	82	2.7	3159	18	US-11-072-512-762		Sequence 762, App
10	60.4	2.0	730	7	US-09-925-065A-5628		Sequence 5628, Ap
11	60.4	2.0	730	11	US-10-301-480-108685		Sequence 108685, App
12	60.4	2.0	730	12	US-10-301-480-720274		Sequence 720274, App
13	55.4	1.8	2217	18	US-11-072-512-201		Sequence 201, App
14	54.6	1.8	1297	18	US-11-096-568A-4896		Sequence 4896, Ap

Db 601 GAAATCATGTATTGTGAAGGAAATGAAAAAATGAAGAACAAAGCGGGAAGAGAGGCG 660  
Qy 1346 CCGACAGGCCGAGATGAAGGAGAGAGAGGTGGCGGTGATGAACACAGGATCCGGGCAATCC 1405  
Db 661 CAGAACTCGGAAATGAAGAAATGAAGAG- - - - -CTCAGGAATATGACAGCAGCTTTCC 714  
Qy 1406 AAATCGGAGAGCGGCGCAATGATACGCGAATATCAGAGACGCTGGAATTTGTCCGCT 1465  
Db 715 AAATCGGAAATTTGCCGATGATTAAGAAATTCGGGTACTATAGAAATGTACCCACT 774  
Qy 1466 GTCGATGCCAGGCGTGCATGACCATCAGATCAAGTGTGCGTGGCAAGCGTCCCAT 1525  
Db 775 TACTCTGACTGATCTACGGAAGAGCAGAGATCTGCGTGTGTGTGAGGAGCGCCACT 834  
Qy 1526 TAGCGCAAGGAGTCAATCGCAAGAGATCGATGTCAATTTCCGTGCGCGCAAGGACAT 1585  
Db 835 GAATGAAGCAAGAACTGGCCCAAGAAAGAAATCGATGTGATTTCTGTTCAGCAAGTGTGT 894  
Qy 1586 GCTCATCTGTCAGAGCGCGCAGCAAGGTGCACTCAACAAAGTTCTCGGAGAACCA 1645  
Db 895 GCTCTTTGTGCATGAACCAATTTGAAGTGAATTTGACAAAGTATCTGGAGATCAGGC 954  
Qy 1646 GTTTCGCTTGACTAGCTTCAACGACACGTCGACCAATGCAATGCTATGATATCAATAC 1705  
Db 955 ATTCTGCTTCGATTTTGCTTGATGAACAAAGCTTCAATGAAGTCTGTACAGGTTTCC 1014  
Qy 1706 AGCAAGCGCTTGTGAAACCAATTTTCGAGGCGGAATGGGAGTCTTTCGCTACGG 1765  
Db 1015 AGCAAGGCGACTGGTACAGAGATTTTGAAGGGGGAAGCAACCTCTTTCGCTATGG 1074  
Qy 1766 CCAGACGGGATCGGGCAAAACCCACACATCGGCGGTGAGTTTAATGAAAGGTGCAGGA 1825  
Db 1075 GCAGACGGGAGTGGGAAGCTCACAAATGGGTGGAGACCTGTCTGTAAATCTCAGAA 1134  
Qy 1826 CTGAAGAAAGGATCTAGCCCATGGCGGCCCAAGGATCTTTGTGACCTTGAATATGCC 1885  
Db 1135 TACATCTAAAGGATCTTGAATGGCTTCCGGGATGCTTCTCTGAGAGGTGAGCC 1194  
Qy 1886 GCGTTACCGGCCCATGATCTAGTCTCGGCCAGAGGATCTTTGTGACCTTGAATATGCC 1945  
Db 1195 TCGCTACCGGAATTTAAACTTTGGAAGTTTATGTGACATTTCTTGAGATCTCAATGGAA 1254  
Qy 1946 GGTCTTCGATCTTCTGTCGCAAGCAAGCAAACTGCGCGTCTCGGAGGATGTAACAGCA 2005  
Db 1255 GGTGTTGATCTGCTCAACAGAGGCCAAGTACTGTATAGAGACACAGCAGCAGCA 1314  
Qy 2006 AGTCAGGTGTGGACTCAACGGAAGGTGGTGCATGGCGTTCGAGGAGGTACTGAAGCT 2065  
Db 1315 GGTGAGGTTGTGGACTTTCAGGAGTACCTGGTTAACTGCTGATGATGTCATCAAGAT 1374  
Qy 2066 CATCCAGCACGGCAATGTCGCCAGACATCGGCCAGAGCTGCGGCACTTCCAATTCGTC 2125  
Db 1375 GCTCAACATGGGAGTGTCTGAGACTTTCGGAACAGACTTTTCGCAACTCCAATTTCTC 1434  
Qy 2126 GCGTTTCGACGCGCTTTTCCAGATTTGCTCGCGCGCAGGAGCTCGAGGAGATCCATGG 2185  
Db 1435 CCGCTCCCATGCTGCTTCCAGATTTCTTTCGAGCCAAAGG- - - - -GAGATTGCATGG 1488  
Qy 2186 CAAGTTCTGCTTCATTCGATCTGGCGGCAATGAGCGGGCGTGGACACTTCTCTCGGCGGA 2245  
Db 1489 CAAATTTCTTCTTGGTGGATCTGGCAGGGAATGAACGAGGGGCTGACACTTCTAGTGCTGA 1548  
Qy 2246 TCGCAGACGGTATGGAGGGTCCGAGATTAACAAATTCGTGTGCGCCCTCAAGGAGTG 2305  
Db 1549 CCGCAGACTCGATGGAGGGTGCAGAGATTAACAAAGTCTCTCTAGCACTGAAGGAGTG 1608  
Qy 2306 CATTCGTCGTTGGGCAAAAGTGGGCCCACTTGCCTTTCGCTGCTCTCAACTCACCA 2365  
Db 1609 CATCAGGCGCTGGGACAGAACAGGCTCACACCCCATTTCCGTGAGAGCAAGCTGACTCA 1668  
Qy 2366 GGTGCTGGCGACTGCTTCATTTGGCGGAGAGAGCAAGCTGTCATGATGATGATGATGATCTC 2425

RESULT 2

US-11-128-049-3433  
; Sequence 3433, Application US/11128049  
; Publication No. US20060010513A1  
; GENERAL INFORMATION:  
; APPLICANT: Melville, Mark W.  
; APPLICANT: Charlebois, Timothy S.  
; APPLICANT: Mounts, William M.  
; APPLICANT: Hann, Louane E.  
; APPLICANT: Sinacore, Martin S.  
; APPLICANT: Leonard, Mark W.  
; APPLICANT: Brown, Eugene L.  
; APPLICANT: Miller, Christopher P.  
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR  
; TITLE OF INVENTION: MAKING AND USING SAME  
; FILE REFERENCE: 01997.027700  
; CURRENT APPLICATION NUMBER: US/11/128.049  
; CURRENT FILING DATE: 2005-05-11  
; PRIOR APPLICATION NUMBER: US 60/570,425  
; PRIOR FILING DATE: 2004-05-11  
; NUMBER OF SEQ ID NOS: 7285  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 3433  
; LENGTH: 2703  
; TYPE: DNA  
; ORGANISM: Cricetus griseus  
US-11-128-049-3433

Query Match 15.0%; Score 461.4; DB 17; Length 2703;  
Best Local Similarity 61.5%; Pred. No. 1.1e-107;  
Matches 779; Conservative 0; Mismatches 476; Indels 12; Gaps 2;

Qy 1226 GGCGGCACAGGAGTCCGCTCGGCGCACAAACCCAGGAGCTGGCGCGCTAGTACCCG 1285  
Db 541 GGAGGACAGAGAGCAATCCATCCACCCGAGCACATCTTCTGAAAACCTGCTCGAG 600  
Qy 1286 GCGATCCACGCAATTTGAAGAGGTGGAGCGCACTGAAGAGAAATCGCGAAGCGACGCGC 1345  
Db 601 GAAATCATGTATTCTGAAGGAATGGAATAATGAAGAACAGCGGGAAGAGAGGCG 660  
Qy 1346 CCGACAGCGCAGATGAAGAGAGAGAGGTGGCGCTGATGAACAGGATCCGGCAATCC 1405  
Db 661 CCAGAACTCGGAAATGAAGATAAAGCGAG- - - - -CTCAGGAATATGACAGCAGCTTTCC 714  
Qy 1406 AAATCGGAGACGCGCAATGATACCGAATATCAGAGCACCTGGAATTTGTGCGCT 1465  
Db 715 AAATCGGAAATTTGCCCGGATGATTAAGAAATTCGGGTACTATAGAAATGTACCCACT 774  
Qy 1466 GCTCGATGGCGAGCGCTCGATGACCATCAGATGTCGCTGGCGCAAGCGTCCCAT 1525  
Db 775 TACTCTGACTGATCTTACGGAAGACACAGATCTGCTGTGTGTGAGGAAGCGCCACT 834  
Qy 1526 TAGCCGCAAGAGGTCAATCGCAAGAGATCGATGTCAATTTCCGTGCGCGCAAGGACAT 1585  
Db 835 GAATAAGCAAGAACTGGCCCAAGAAAGAAATCGATGTGATTTCTGTTCCAGCAAGTGT 894  
Qy 1586 GCTCATCTGTCAGAGCGCGCAGAGAGGTGCACTCACCAAGTTCTCTGAGAGACCA 1645  
Db 895 GCTTTTGTGATGAACCCAAATTTGAAGTGGACTTGCAAAAGTATCTGGAGATCAGGC 954  
Qy 1646 GTTTCGCTTCGATGAGCTTCAACGACACGTCGCAATGCCATGCTGATATCAAAATAC 1705



Db 955 ATTCTGCTTCGATTTTGCCTTTGATGAAACAGCTTCAATGAAGTGTCTACAGGTTCCAC 1014  
Qy 1706 AGCCAGCGCTGTGTGAAAACCAATTTTCGAGGGCGAATCGCGACGTCTTCGCTACCG 1765  
Db 1015 AGCAAGCCACTGTGTACAGACGATTTTGAAGGGGAAAGCAACCTGCTTTGCCTATGG 1074  
Qy 1766 CCAGAGGGATCGGCGAAAGGACACCATGCGGGGTGAGTTTAATGGAAGGTGCGAGGA 1825  
Db 1075 GCAGAGGGCAGTGGGAAGACTCACAAATGGGTGGAGACCTGTCTGTGTAATCTCGAA 1134  
Qy 1826 CTGCAAGAAACGGCATCTACGCCATGCGCGCAAGGATGCTTTCTGTGACCCCTGAATATGCC 1885  
Db 1135 TACATCTAAGGGATCTATGCAATGGCTTCCGGGATGTCTTCTCTGAAGAGTCAAGC 1194  
Qy 1886 GCGTTACCGCGCCATGAATCTAGTCGTCTCGGCCAGTTTCTTTGAGATTTACAGTGGCAA 1945  
Db 1195 TCGCTACCGGAACTTAACTTGAAGTTTATGTGACATTTCTTTGAGATCTACAATGGAA 1254  
Qy 1946 GGTCTTCGATCTTCTGTCGACAGCAGAACTGCGCGTCTCGAGGATGTTAAACAGCA 2005  
Db 1255 GGTGTTTGATCTGCTCAACAGAGGCAAGCTACGTTACTAGAAGACAGCAAGCAGCA 1314  
Qy 2006 AGTCGAGGTGGTGGGACTCACCGAGAAAGTGTGTCGATGGCGTCTCGAGGATCTGAAGCT 2065  
Db 1315 GGTGCGAGTTGTGGACTTCAGAGTACCTGTTAACTGTGCTGATGATGTCAAGAT 1374  
Qy 2066 CATCCAGCAGCGCAATGCTGCCCGCAACATCCGCCAGACGTCCGCCCAACTCCAAATTCGTC 2125  
Db 1375 GCTCAACATGGGCGTCTGAGGACTTCTGGACAGACTTTTGGCCAACTCCAAATCTTTC 1434  
Qy 2126 GGTTCGACCGCGTTTTCAGATTTGTGTCGGCGCGCAGGGCTCGACGAAAGATCCATGG 2185  
Db 1435 CCCTCCCATGCTTCCAGATTTCTTCTTCGAGCCAAAGG-----GAGATTGATGG 1488  
Qy 2186 CAAGTTCTGTCATCGATCTGCGGCAATGAGCGGGGTGAGACACTTCTTCGCGCGA 2245  
Db 1489 CAATTTCTCTGCTGATCTGCGAGGAATGAACGAGGGCTGACACTTCTAGTCTGA 1548  
Qy 2246 TCGCGAGACGCTATGAGGGTCCGAGATTAACAAATCGCTGCTCGGCCCTCAAGGAGTG 2305  
Db 1549 CGCCAGACTCGCATGGAGGTGACAGAGTTAAACAGTCTCTCTAGCAGTGAAGGATG 1608  
Qy 2306 CATTCGTGCTGGGCAACAGTTCGGCCACTTGGCCCTTCGCTGTCTCCAACTCAACCA 2365  
Db 1609 CATCAGGGGCTGGGACAGAAAGGCTCACACCCCAATTCCTGAGCAGCAAGCTGACTCA 1668  
Qy 2366 GGTGCTGCGCGACTCGTTCAATGCGGAGAGACAGAGCTGATGATAGCCATGATCTC 2425  
Db 1669 GGTGCTAAGGATTCATTCATGGGAGAACTCGAGGACTTGCATGATTTGCCATGATCTC 1728  
Qy 2426 GCGGGACTTTAGCTCTCGAGCACACGCTCAACAGCTGCGCTATGCGGATCGTGTCAA 2485  
Db 1729 ACCAGGATTAAGTCTCTGTGAATATCTTTAAACAGCTGAGATATGACAGACAGTCAA 1788  
Qy 2486 GGAGCTG 2492  
Db 1789 GGAGCTG 1795  
|||||

## RESULT 3

US-11-145-307A-22  
; Sequence 22, Application US/11145307A  
; Publication No. US20060094035A1  
; GENERAL INFORMATION:  
; APPLICANT: Acturus Bioscience, Inc.  
; APPLICANT: Erlander, Mark G.  
; APPLICANT: Ma, Xiao-Jun  
; TITLE OF INVENTION: Identification of Tumors  
; FILE REFERENCE: 022041-002020US  
; CURRENT APPLICATION NUMBER: US/11/145,307A  
; CURRENT FILING DATE: 2005-06-03  
; PRIOR APPLICATION NUMBER: US 60/577,084  
; PRIOR FILING DATE: 2004-06-04

; NUMBER OF SEQ ID NOS: 268  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 22  
; LENGTH: 2832  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-145-307A-22

Query Match 14.7%; Score 453; DB 14; Length 2832;  
Best Local Similarity 62.5%; Pred. No. 1.6e-105;  
Matches 746; Conservative 0; Mismatches 435; Indels 12; Gaps 2;

Qy 1300 TGAAGAGGTGGAGCGACTGAAGGAGAAATCGCGAGAACGACGCGCCGACAGGCCGAGA 1359  
Db 647 TGAAGGAGTGGAAAAAATGAAGAAACGAGCGAGAGAGAAAGAGGCCAGCACTCTGAAA 706  
Qy 1360 TGAAGGAGGAGAGGTGGCGCTGATGAACAGGATCCGGGCAATCCAAACTGGGAGACGG 1419  
Db 707 TGAAGATGAAGAGAG-----CTCAGGAGTATACAGTAGTTTTTCCAACTGGGAATTG 760  
Qy 1420 CGCAATGATACCGCAATATCAGAGCAGCTCGAAATTTGTGCGGCTGCTCGATGGCCAGG 1479  
Db 761 CCCGATGATTAAGAAATTTCCGGCTACTTTGGAATGTATCTCACTTACTATGACTATC 820  
Qy 1480 CCGTCGATGACCATCAGATCAGATGTGTGTCGCGAAGCGTCCCATTTAGCCGCAAGAGG 1539  
Db 821 CTATCGAAGAGCACAGATATGTGTCTGTGTTAGGAAACGCCCACTGAATGAAGCAAGAT 880  
Qy 1540 TCATCGCAGGAGATCGATGTCTATTTCCGTGCGCGCAGGACATGCTCATCTGTGACG 1599  
Db 881 TGCCCAAGAAAGAAATTTGATGTGATTTCCATTCCTAGCAAGTGTCTCTCTTGGTACATG 940  
Qy 1600 AGCGCGCAGCAGAGGTGCGACCTCACCAAGTTCTCTGAGAAACCAACAAATTTTCGCTTCGACT 1659  
Db 941 AACCCTAAGTTGAAGTGGACTTAAACAAAGTATCTGGAGAAACCAAGCATTTCTGCTTACT 1000  
Qy 1660 AGCCTTCAACGACAGCTGCGACCAATGCCATGTGTATACAAATACACAGCAAGCCGTGG 1719  
Db 1001 TTGCATTTGATGAAACAGCTTCGAATGAAGTTGTCTACAGGTTTACAGCAGCAAGCCACTGG 1060  
Qy 1720 TGAACACCAATTTTCGAGGGCGGAAATGGCGACGTGCTTCGCTTACGCGCAGCAGCGGATCGG 1779  
Db 1061 TCAGACCAATCTTTGAAGGTGGAAAGCAACTTGTGTTGATATGGCCAGCAGCAAGTGTG 1120  
Qy 1780 GCAAAACGACACCACTGCGCGGTGAGTTTAAATGGAAGGTGACAGGACTGCAAGAACCGCA 1839  
Db 1121 GCAGACACATATCTATGGCGGAGACTCTCTGGGAAAGCCAGATGCAATGCCAAGGGA 1180  
Qy 1840 TCTACGCCATGCGCGCAAGGATGTCTTTGTGACCTGTAATATGCGCGGTACCGCGCCA 1899  
Db 1181 TCTATGCCATGGCTCCCGGACGCTTCTCTCTGAAGAAATCAACCTGTCTACCGGAAGT 1240  
Qy 1900 TGAATCTAGTCTCTCGCGCAGTCTTTTGTGATTTTACAGTTCGACAGGTTCTTCATCTTC 1959  
Db 1241 TGGGCTTGAAGTCTATGTGACATTTCTCGAGATCTACAATGGGAAGCTGTTTGAAGTCTG 1300  
Qy 1960 TGTCCGACAAAGCAAGAACTGCGCGCTCTGAGGATGTTAAACAGCAAGTGCAGGTGTGG 2019  
Db 1301 TCACAAAGAAAGCCAGCTGCGCGTCTGAGGAGCGGCAAGCAAGTGCAGTGTGG 1360  
Qy 2020 GACTCACCGAGAGGTGTGTCGATGCGTCGAGAGGATCTGAAGCTCATCCAGCAGGCA 2079  
Db 1361 GGTGCGAGGACATCTGTTAACTCTGCTGATGATGTCTCATCAAGATGATCGACATGGCA 1420  
Qy 2080 ATGCTGCCGCAACATCCGCGCAGACGTGCGCCCACTCCAAATTCGTGCGTTCGACGCGC 2139  
Db 1421 GCGCCTGCGAGAACCTCTGGGCGAGACATTTGCCAACTCCAAATTCCTCCCGCTCCACGCGT 1480  
Qy 2140 TTTTCCAGATTTGTGCTGCGCGCGCAGGGCTCGACGAAAGATCCATGCAAGTTCTTCGTTC 2199  
Db 1481 GCTTCCAAATTAATTTCTTCGAGCTAAGG-----GAGATGCAATGCAAGTCTCTTTGG 1534  
Qy 2200 TCGATCTGCGGGCAATGAGCGGGCGTGGACACATCTTCTTCGCGCGATCGGACAGCAGCA 2259

Db 1535 TAGATCTGCGAGGAATGAGCGGCGGACACTTCCAGTGTGACCGGCAGACCGCA 1594  
Qy 2260 TGGAGGGTCCGAGATTAAACAAATCGCTGCGCCCTCAAGAGTGCATTCTGTGTTGG 2319  
Db 1595 TGGAGGGCCGAGAAATCAACAAGTCTCTTGAAGGCTGAAGAGTGCATCAGGGCCCTGG 1654  
Qy 2320 GCAACAGTCCGCCCACTTGCCTTCCCGTGTCTCCAAACTCACCCAGGTGCTGCGGACT 2379  
Db 1655 GACAGAACAAAGCTCACACCCGTTCCGTGAGAGCAAGCTGACAGGTGCTGAGGACT 1714  
Qy 2380 CTTTCAATGGCGAGAGAGCAAGACGTGCATGATAGCCATGATCTGCGCGGACTTAGCT 2439  
Db 1715 CTTTCAATGGCGAGAACTCTAGGACTTGCATGATGCGCACAGTCTCACAGGCATAAGCT 1774  
Qy 2440 CTTGCGGACACAGCTCAACAGCTGCGGTATGCGGATCGTGTCAAGAGCTG 2492  
Db 1775 CCTGTGAATATACTTTAAACACCCCTGAGATATGACAGACAGGGTCAAGAGCTG 1827

## RESULT 4

US-10-955-054A-137  
; Sequence 137, Application US/10955054A  
; Publication No. US20050266420A1  
; GENERAL INFORMATION:  
; APPLICANT: PUSZTAI, LAJOS  
; APPLICANT: SYMMANS, W. FRASER  
; APPLICANT: HESS, KENNETH R.  
; APPLICANT: AYERS, MARK  
; APPLICANT: STEC, JAMES  
; TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY  
; FILE REFERENCE: UTXC:880US  
; CURRENT APPLICATION NUMBER: US/10/955,054A  
; CURRENT FILING DATE: 2004-09-30  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 137  
; LENGTH: 2401  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-955-054A-137

Query Match 14.6%; Score 449.8; DB 10; Length 2401;  
Best Local Similarity 62.4%; Pred. No. 1e-104;  
Matches 744; Conservative 0; Mismatches 437; Indels 12; Gaps 2;  
Qy 1300 TGAAGAGGTGGAGCGACTGAAGAGAAATCGGAGAACGCGCGCCCGACAGGCCGAGA 1359  
Db 770 TGAAGGAAGTGGAAAAATGAAGAAACAAGCGAGAGAGAAAGGCCCGCAACTCTGAAA 829  
Qy 1360 TGAAGGAGGAGAGGTGGCGCTGATGAACACGAGATCCGGGCAATCCAACTGGGAGACGG 1419  
Db 830 TGAGAAATGAAGAG-CTCAGGAGTATGACAGTAGTCTTCCAACTGGGAAATTTG 883  
Qy 1420 CGCAAAATGATACGGGAATATACAGACAGCTGGAAATTTGTGCGCTGCTCGATGGCCAGG 1479  
Db 884 CCGGAATGATTAAGAAATTTCCGGCTACTTTGGATGTTCATCCACTTACTATGACTGATC 943  
Qy 1480 CCGTGTGATGACATCAGATCAGTGTGCGTGGCGGACGCTGCCATTAAGCCGCAAGGAGG 1539  
Db 944 CTATCGAAGAGCAGACAGATATGTGTGTGTAGGAAACGCGCCACTGAATAAGCAAGAAAT 1003  
Qy 1540 TCAATCGCAAGGAGATCGATGCTATTTCCGTGCGCGCAGACAGCATGCTATCGTCAAG 1599  
Db 1004 TGGCAAGAAAGAAATGATGTGATTTCCATTCCTAGCAAGTGTCTCTCTTGTGTACATG 1063  
Qy 1600 AGCCGCGCAGCAAGTTCGACCTCACCAAGTTCTTGGAGAACCAAGTTTCGCTTCGACT 1659  
Db 1064 AACCACAGTTGAAAGTGGACTTAACAAAGTATCTGGAGAACCAAGCACTTCTGCTTGAAT 1123  
Qy 1660 ACGCTTCAACGACACGTCGACATGCCATGCTGATATACAAATACACAGCCAGCGTTGG 1719  
Db 1124 TTGCATTTGATGAACACAGCTTTCGAATGAAGTTGTCTACAGGTTTCACAGCAAGGCCACTGG 1183

Qy 1720 TGAAAACCATTTTCGAGGCGGGAATGGCGAGCTGTGCTTCCCTACGGCCAGACGGGATCGG 1779  
Db 1184 TACAGACAAATCTTTGAAGGTGAAAGCAACTTGTTTTTCATATATGGCCAGACAGGAAGTG 1243  
Qy 1780 GCAAAAACGACACACCATCGGCGGTGAGTTTAATGAAAAGGTGCAGGACTGCAAGAACGGCA 1839  
Db 1244 GCAAGACACATACTATATGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCAATCCAAAGGGA 1303  
Qy 1840 TCTAGCCATCGGCGGCAAGGATGTCTTTTGTGACCCCTGAATATGCGCGTTTACCGCGCA 1899  
Db 1304 TCTATGCCATGGCTTCCCGGACGCTTCTCTCTGAAGAAATCAACCTCTGTCTACCGGAAAT 1363  
Qy 1900 TGAATCTAGTCTGTCTCGCCAGTCTTCTTGAATTTTACAGTGGCAAGTCTTCGATCTTC 1959  
Db 1364 TGGGCTTGAAGTCTATGTGACATCTTCGAGATCTCAATGGGAAGCTGTTTGACCTGTC 1423  
Qy 1960 TGTCCGACAAACGACAGAAACTGCGCGCTCTCGAGGATGGTAAACAGCAAGTGCAGGTGGTGG 2019  
Db 1424 TCAACAGAAAGCCAGCTGCGCGTCTGGAGGACGCGCAACAGACAGTGCAGTGGTGG 1483  
Qy 2020 GACTCACCGAAGAGTGTGATGCGGTGCGAGGAGTACTGAAGTCAATCAGCACGGCA 2079  
Db 1484 GGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCAATCAAGATGCTCGACATGGGCA 1543  
Qy 2080 ATGCTGCCCGCAACATCGGCCAGAGCTCGGCCAACTCCAATTCGTCGCTTCGCAACGCG 2139  
Db 1544 CGGCTGCGAGAACTCTGGGCGAGACATTTGCAACTCCAATTCCTCCGCTCCCAACGCGT 1603  
Qy 2140 TTTTCCAGATTGTCTGCGCGCGCAGGCTCGACAAAGATCCATGGGCAAGTTCTCGTTCA 2199  
Db 1604 GCTTCCAAATTTTCTTCGAGCTAAAGG-----GAGATGCAATGGCAAGTTCTCTTTGG 1657  
Qy 2200 TCGATCTGGCGGCAATGAGCGGGCGGTGGACACTTCTCTCGGCCGATCGGCGAGCGGTA 2259  
Db 1658 TAGATCTGGCAGGAATGAGCGAGCGCAGACACTTTCAGTGTCTGACCGGCGAGACCCGCA 1717  
Qy 2260 TGGAGGGTGGCGAGATTAAACAAATCGTGTGCGCCCTCAAGGAGTGCATTCGTGCTGG 2319  
Db 1718 TGGAGGGCGGAGAAATCAACAAAGATCTCTTAGGCCCCCTGAAGGAGTGCATCAGGGCCCC 1777  
Qy 2320 GCAAAACAGTTCGGCCCACTTGCCTTCCGTGTCTCCAAACTCACCCAGGTGCTGCGCGACT 2379  
Db 1778 GACAGAACAGGCTCACACCCCGTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGACT 1837  
Qy 2380 CGTTTATTTGGCGAAGAGCAAGACGTGTCATGATAGCCATGATCTCGCCGGGACTTAGCT 2439  
Db 1838 CTTTCAATTTGGGGAGAACTTAGGACTTTCATGATTTGCCATGATCTCACAGGCAATAAGCT 1897  
Qy 2440 CTTGCGAGCACACCTCAACACGCTGCGCTATGCGGATCGTGTCAAGGAGCTG 2492  
Db 1898 CCTGTGAATATACTTTTAAACACCCCTGAGATATGACAGACAGGGTCAAGAGCTG 1950

## RESULT 5

US-10-955-054A-101  
; Sequence 101, Application US/10955054A  
; Publication No. US20050266420A1  
; GENERAL INFORMATION:  
; APPLICANT: PUSZTAI, LAJOS  
; APPLICANT: SYMMANS, W. FRASER  
; APPLICANT: HESS, KENNETH R.  
; APPLICANT: AYERS, MARK  
; APPLICANT: STEC, JAMES  
; TITLE OF INVENTION: MULTIGENE PREDICTORS OF RESPONSE TO CHEMOTHERAPY  
; FILE REFERENCE: UTXC:880US  
; CURRENT APPLICATION NUMBER: US/10/955,054A  
; CURRENT FILING DATE: 2004-09-30  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 101  
; LENGTH: 2740  
; TYPE: DNA

; ORGANISM: Homo sapiens  
US-10-955-054A-101

Query Match 14.6%; Score 449.8; DB 10; Length 2740;  
Best Local Similarity 62.4%; Pred. No. 1e-104;  
Matches 744; Conservative 0; Mismatches 437; Indels 12; Gaps 2;

```
QY 1300 TGAAGAGGTGGAGGACTGAAGAGAGATCGGAGAGCGACGCGCGGACAGCGCGGAGA 1359
DB 638 TGAAGAGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 697
QY 1360 TGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1419
DB 698 TGAGAGTGAAGAGG-----CTCAGGAGTATGACAGTAGTCTTCCAACTGGGAATTG 751
QY 1420 CGCAATGATACGCGAATATACAGACACGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1479
DB 752 CCGCAATGATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 811
QY 1480 CCGTGCATGACCATCAGATCAGATCAGATGCTGCGGAGGAGGAGGAGGAGGAGGAGG 1539
DB 812 CTATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 871
QY 1540 TCAATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1599
DB 872 TGCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 931
QY 1600 AGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1659
DB 932 AACCCCAAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 991
QY 1660 AGCGCTTCAACGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1719
DB 992 TTGCATTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1051
QY 1720 TGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1779
DB 1052 TACAGAGCAATCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1111
QY 1780 GCAAAACGACACCATGCGGCGGTGAGTTTAAATGGAAGGTGACAGGAGTGCAGAAACGGCA 1839
DB 1112 GCAAGACACATACTATGGGCGGAGACTCTCTCGGAAAGGCCAGAGATGCATCCAAAGGGA 1171
QY 1840 TCTACGCCATGGCGGCGCAAGGATGCTTTGTGACCTGCAATATGCGCGGTACCGCGCCA 1899
DB 1172 TCTATGCCATGGCCTCCGCGAGCTCTTCCTCTGGAAGATCAACCTGTCTACCGGAAGT 1231
QY 1900 TGAATCTAGTCTGCGCGCAGTTCTTTGAGATTTACAGTGGCAGAGTCTTCGATCTTC 1959
DB 1232 TGCGCCTGGAAGTCTATGTGACATCTTCGAGATCTACAAATGGGAAGCTGTTTGACCTGC 1291
QY 1960 TGTCCGACAGCAGAAACTGCGCGTCTCGAGGAGTGGTAAACAGCAAGTGCAAGTGGTGG 2019
DB 1292 TCAACAAGAGGCGCAAGCTGCGGTCTGGAGGACGCGCAAGCAAGAGTGCAAGTGGTGG 1351
QY 2020 GACTACCGAGAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2079
DB 1352 GGCTGCGAGGAGCATCTGGTTAACTCTGCTGATGATGATCATCAGATGCTCGACATGGGCA 1411
QY 2080 ATGCTGCCGACATCCGCGCAGAGCTCGCGCAACTCCAAATTCGTGCGGTTCGACGCGG 2139
DB 1412 GCGCCTGCAAGACCTCTGGGAGACATTTGCCAACTCCAAATTCCTCCCGCTCCACGCGT 1471
QY 2140 TTTTCCAGATTTGCTGCGCGCGCAGGCTCGACGAGAGTCCATGGCAAGTCTCGTTCA 2199
DB 1472 GCTTCCAAATTTCTTCGAGCTAAAG-----GAGATGATGATGCAAGTCTCTTTGG 1525
QY 2200 TCGATTCGCGGCAATGAGCGGGCGTGGACATTTCTCGCGCGAGTCCGCGAGACCGGTA 2259
DB 1526 TAGATTCGCGGAGGAGTGAAGCGAGGCGGAGACACTTCCAGTGTGAGCGCGGAGACCGCA 1585
QY 2260 TGGAGGTCGCGAGATTAACAAATCGCTGCTGCGCCTCAAGGAGTGCATTCGTGCGTTGG 2319
```

```
DB 1586 TGGAGGCGCGAGAAATCAACAGAGTCTCTTATGCCCTGAAGGAGTGCATCAGGCGCTCG 1645
QY 2320 GCAAAACAGTCGCGCCCACTTGGCCCTTCGGTGTCTCCAACTCACCAGGTCCTGCGCGACT 2379
DB 1646 GACAGAAACAGGCTCACACCCCGTTCGTTGAGAGCAAGCTGACACAGTCTCTGAGGACT 1705
QY 2380 CGTTCAATTGGCGAGAGAGCAAGCTGCTGATGATAGCATGATCTGCCGCGGACTTACGT 2439
DB 1706 CTTTCAATTGGGAGAACTCTAGGACTTGCATGATTCACGATCTCACAGGATTAAGCT 1765
QY 2440 CTTGCGAGCAGCAGCTCAACACGCTGCTGCTGATCGGTGCTGTCAGGAGCTG 2492
DB 1766 CTTGCTGATATATCTTTAAACACCTGAGATATGACAGAGGCTCAAGAGCTG 1818
```

## RESULT 6

US-10-960-414-299  
; Sequence 299, Application US/10960414  
; Publication No. US20060074565A1  
; GENERAL INFORMATION:  
; APPLICANT: MILLER, LANCE D.  
; APPLICANT: GEORGE, JOSH  
; APPLICANT: VEGA, VINSENSIUS B.  
; TITLE OF INVENTION: METHODS, SYSTEMS, AND COMPOSITIONS FOR CLASSIFICATION,  
; TITLE OF INVENTION: PROGNOSIS, AND DIAGNOSIS OF CANCERS  
; FILE REFERENCE: 38271-76067  
; CURRENT APPLICATION NUMBER: US/10/960,414  
; NUMBER OF SEQ ID NOS: 500  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 299  
; LENGTH: 2740  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-960-414-299

Query Match 14.6%; Score 449.8; DB 13; Length 2740;  
Best Local Similarity 62.4%; Pred. No. 1e-104;  
Matches 744; Conservative 0; Mismatches 437; Indels 12; Gaps 2;

```
QY 1300 TGAAGAGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1359
DB 638 TGAAGAGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 697
QY 1360 TGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1419
DB 698 TGAGAGTGAAGAGG-----CTCAGGAGTATGACAGTAGTCTTCCAACTGGGAATTG 751
QY 1420 CGCAATGATACGCGAATATCAGAGCACGCTGGAATTTGTGCGGTGCTGCTGAGTGGCGAGG 1479
DB 752 CCGCAATGATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 811
QY 1480 CCGTGCATGACCATCAGATCAGATGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1539
DB 812 CTATCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 871
QY 1540 TCAATCGCAAGGAGATCGATGCTTTCGGTGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1599
DB 872 TGCGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 931
QY 1600 AGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1659
DB 932 AACCCCAAGTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 991
QY 1660 AGCGCTTCAACGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1719
DB 992 TTGCATTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1051
QY 1720 TGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1779
DB 1052 TACAGAGCAATCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1111
QY 1780 GCAAAACGACACCATGCGGCGGTGAGTTTAAATGGAAGGTGCGAGGAGTGCAGAAACGGCA 1839
```

Db 1112 GCAAGACACATCTATGCGGAGACCTCTCTGGGAAAGCCAGAAATGATCCAAAGGA 1171  
Qy 1840 TCTAGCCATGGGGCAAGGATGCTTTGTGACCTGAAATATGCGCGTTACCGGCA 1899  
Db 1172 TCTATGCCATGGCTCCCGGAGCTCTTCTCTGGAAGATCAACCCCTGCTACCGGAAGT 1231  
Qy 1900 TGAATCTAGTCTCTCGGCCAGTCTTTTGGAGATTTACAGTGGCAGGCTCTCGATCTTC 1959  
Db 1232 TGGGCTGGAAGTCTATGTGACATCTTCGAGATCTACATGGGAGCTGTTTGACCTGC 1291  
Qy 1960 TGTCCGACAAAGCAGAAATGCGCTCTGAGGATGTTAAACAGCAAGTGCAGGTGGTGG 2019  
Db 1292 TCAACAAGAAGCCAAAGCTGCGCTGCTGGAGACGGCAAGCAACAGGTGCAAGTGTGG 1351  
Qy 2020 GACTCACCGAGAAGTGTGATGAGTGGCTGAGAGGTACTGAAGCTCATCGACAGCGCA 2079  
Db 1352 GGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCAATCAAGATGCTCGACATGGCA 1411  
Qy 2080 ATGCTGCCCGAACATCCGGCCAGACGTCGGCCAACTCCAAATTCGTCGGTTTCGACGCG 2139  
Db 1412 GCGCTGAGAACCTCTGGGAGACATTTGCCAATCTCCAAATCTCCCGCTCCCAAGCT 1471  
Qy 2140 TTTTCCAGATGCTGCGGCCCGCAGGCTCGACGAAGATCCATGGCAAGTTCTCGTTCA 2199  
Db 1472 GCTTCCAAATATTCTTGAGCTAAAG-----GAGATGATGGCAAGTTCTCTTGG 1525  
Qy 2200 TGATCTGGCGGCAATGAGCGGGGCTGGACACTTCTCGGCCCATCGGAGAGCGGTA 2259  
Db 1526 TAGATCTGGCAGGGAATGAGCGAGCGCAGACACTTCCAGTCTGACCGGAGACCCGCA 1585  
Qy 2260 TGGAGGCTGCCAGATTAACAAATCGCTGCGCCCTCAAGAGTGCAATCGTGGCTGG 2319  
Db 1586 TGGAGGCGCAGAAATCAACAGATCTCTTAGCCCTGAAGAGTGCATCAGGCGCCCTGG 1645  
Qy 2320 GAAACAGTCCGCCCACTTGCCTTCCTGCTTCCAAACTCACCAGAGTGTGCGCCACT 2379  
Db 1646 GACGAACAAGGCTCACACCCGTTCCGTGAGAGCAAGCTGACACAGTGTGAGGACT 1705  
Qy 2380 GCTTCATTTGGGAGAGCAAGAGCTGATGATGAGATGATCTCGCGGGAATAGCT 2439  
Db 1706 CTTTCATTTGGGAGAACTCTAGGACTTGCATGATTTGCCACGATCTCACCGGCATAAGCT 1765  
Qy 2440 CTGGGAGCACAGCTCAACAGCTGCGCTATGCGGATCGTCAAGGAGCTG 2492  
Db 1766 CTTGGAATATCTTTAAACCCCTGAGATATGCACAGGCTCAAGGAGCTG 1818

RESULT 7  
US-11-128-061-7075  
; Sequence 7075, Application US/11128061  
; Publication No. US20060003958A1  
; GENERAL INFORMATION:  
; APPLICANT: Melville, Mark W.  
; APPLICANT: Charlebois, Timothy S.  
; APPLICANT: Mounts, William M.  
; APPLICANT: Hann, Louane E.  
; APPLICANT: Sinacore, Martin S.  
; APPLICANT: Leonard, Mark W.  
; APPLICANT: Brown, Eugene L.  
; APPLICANT: Miller, Christopher P.  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES RELATED TO OLIGONUCLEOTIDE ARRAYS  
; FILE REFERENCE: 01997.027701  
; CURRENT APPLICATION NUMBER: US/11/128,061  
; PRIOR FILING DATE: 2005-05-11  
; PRIOR APPLICATION NUMBER: US 60/570,425  
; PRIOR FILING DATE: 2004-05-11  
; NUMBER OF SEQ ID NOS: 7285  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 7075  
; LENGTH: 1400  
; TYPE: DNA

; ORGANISM: Cricetulus griseus  
US-11-128-061-7075  
Query Match 7.9%; Score 243; DB 17; Length 1400;  
Best Local Similarity 69.6%; Pred. No. 9e-52;  
Matches 346; Conservative 0; Mismatches 145; Indels 6; Gaps 1;  
Qy 1996 GTAAACACAGTAGTCAGGTGGGACTCACCGAGAGAGGTGGTGCATGCGCTCGAGGAGG 2055  
Db 2 GCAAGCAGCAGGTGCGAGGTGTGGGACTTCAGGAGTACCTGCTTAACTGTGCTGATGATG 61  
Qy 2056 TACTGAAAGCTCATCCAGCACGCAATGCTGCCCAACATCCGGCAGACGTCGGCCAACT 2115  
Db 62 TCATCAAGATGCTCAACATGGGCACTGCTCGAGACTTCTTGGACAGACTTTTGGCAACT 121  
Qy 2116 COAATTCGTCGCTTCGCAACGCGCTTTTCCAGATTTGTCTCGGCCGCGAGGGCTCGACGA 2175  
Db 122 CCAATTTCTCCGCTCCCATGCTTCCAGATTCTTCTTCGAGCCAAAGG-----GA 175  
Qy 2176 AGATCCATGGCAAGTTCCTGTTTCATGCTATGCGGGGCAATGAGCGGGGCTGACACTT 2235  
Db 176 GATTGCATGGCAAAATTCCTTTGGTGGATCTGGCAGGGAATGAACGAGGGCTGACACTT 235  
Qy 2236 CCTCGGCGATCGGCAGACGCGTATGAGGGGTGCCGAGATTAACAAAATCGCTGCTGGGCC 2295  
Db 236 CTAGTGTGACCGCAGACTCGCATGGAGGTGAGAGATTAAAGTCTCTCTAGACAC 295  
Qy 2296 TCAAGGAGTGCATTCGTGCTTGGGCAAAAGTCCGCCCACTTGCCTTCGCTGTCTCCA 2355  
Db 296 TGAAGGAGTGCATCAGGCGCTGGGACAGAAAGGCTCACACCCATTCCGTGAGAGCA 355  
Qy 2356 AACTCACCAGGTGCTGGGAGCTCGTTTCATTGGGAGAGAGCAAGACGTGCTATGATAG 2415  
Db 356 AGCTGACTCAGGTGCTTAAGGGATTCATTTCATTGGGAGAACTCGAGGACTTGCATGTTG 415  
Qy 2416 CCATGATCTCGCGGGACTTAGCTCCTGCGAGCACAGCTCAACACGCTGCGCTATGCGG 2475  
Db 416 CCATGATCTCACCAGGCATAGCTCCTGTGATATATCTTTAAACACGCTGAGATATGCG 475  
Qy 2476 ATCGTGTCAAGGAGCTG 2492  
Db 476 ACAGAGTCAAGGAGCTG 492

RESULT 8  
US-11-128-049-7075  
; Sequence 7075, Application US/11128049  
; Publication No. US20060010513A1  
; GENERAL INFORMATION:  
; APPLICANT: Melville, Mark W.  
; APPLICANT: Charlebois, Timothy S.  
; APPLICANT: Mounts, William M.  
; APPLICANT: Hann, Louane E.  
; APPLICANT: Sinacore, Martin S.  
; APPLICANT: Leonard, Mark W.  
; APPLICANT: Brown, Eugene L.  
; APPLICANT: Miller, Christopher P.  
; TITLE OF INVENTION: OLIGONUCLEOTIDE ARRAYS TO MONITOR GENE EXPRESSION AND METHODS FOR  
; FILE REFERENCE: 01997.027700  
; CURRENT APPLICATION NUMBER: US/11/128,049  
; CURRENT FILING DATE: 2005-05-11  
; PRIOR APPLICATION NUMBER: US 60/570,425  
; PRIOR FILING DATE: 2004-05-11  
; NUMBER OF SEQ ID NOS: 7285  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 7075  
; LENGTH: 1400  
; TYPE: DNA  
; ORGANISM: Cricetulus griseus  
US-11-128-049-7075  
Query Match 7.9%; Score 243; DB 17; Length 1400;





```
; APPLICANT: SEKI, NAOHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOFUKI
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; FILE OF INVENTION: Novel full length cDNA
; FILE REFERENCE: 084335-0191
; CURRENT APPLICATION NUMBER: US/11/072,512
; CURRENT FILING DATE: 2005-03-07
; PRIOR APPLICATION NUMBER: US 60/350,978
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 201
; LENGTH: 2217
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-072-512-201

Query Match      1.8%; Score 55.4; DB 18; Length 2217;
Best Local Similarity 49.0%; Pred. No. 0.0012;
Matches 270; Conservative 0; Mismatches 266; Indels 15; Gaps 4;

QY 1947 GTCTTCGATCTTCTGTCGGAACGAGAACTCGCGTCTCTGGAGGATGTTAAACAGCAA 2006
    |||||
Db 1477 GGCATCTATGTTTCAGACCTCTCAACGACCTCTTCCTGTCCTCATCGAGGAGCCAGCAATGAC 1418

QY 2007 GTCCAGGTGTGGGACTCCCGAGAGAGGTGTCGTGGGTGCGAGGAGTACTGAAGCTC 2066
    |||||
Db 1417 ATGGAGTATGAGGTCTCCATGTCTTACCTGGAGATCATGACGTGCTGATGAAGGGGAAAC 1358

QY 2067 ATCCAGCAGCGCAATGTCGCCGAACATCCGGCCAGACGTCGGCCCACTCCAATTCGTCG 2126
    |||||
Db 1357 CGGCAGAGGACCCAGAGGCCACGGCGGCCAACACGACGTCCTCCGCTCCACGGCGGTA 1298

QY 2127 CGTTCGACGCCGTTTTCAGATTTGTCGGGCCCGCAGGGGTCGACGAAGAT---CCAT 2183
    |||||
Db 1297 CTGCAGGTGACCTGCGGCCAGCGCAGCGCGGTCAAGAACATCTTGCAGGAGGTGCGGCAG 1238

QY 2184 GGCAGTTCTCGTTTCATCATCTGCGGGCAATGAGCGGGCGGTGACACATTCCTCGGCC 2243
    |||||
Db 1237 GCGCGCTGTTCATGATGACCTGCTGGCTCAGAGCGCGCTCGCAGA---CACAGAA 1181

QY 2244 GATCGCAGACGCGTATGAGGGTGCAGAGATTAAACAAATCGCTGCTGCCCTCAAGGAG 2303
    |||||
Db 1180 CGTGGCAGCGTATGAGGAGGGGGCCACATCAACCGCTCACTGCTGCACTGGGCAAC 1121

QY 2304 TGCAATTCGTGCGTTGGGCAACAGTCGCGCCCACTTGCCCTTC-----CGTGTCTCCAA 2357
    |||||
Db 1120 TGCAATCAACGCCCTGAGCGACAAGGGTAGCAACAGTACATCAACTATCGCGACAGCAAG 1061

QY 2358 CTCACCCAGGTGTCGGCAGCTGTTTCATGTCGGAGAGAGCAGCAGCTGATGATAGCC 2417
    |||||
Db 1060 CTCACCCGCTCTCTGAAGACTC---TCTGGGAGGAAACAGCGCAGTGTATGCTGCT 1004

QY 2418 ATGATCTCCCGGGACTTAGCTCTCGCAGACACACGCTCAACACGCTGGCTATGCGGAT 2477
    |||||
Db 1003 CACATCAGTCTCTCGAGCAGTGCCTTCGAGGAGTCCCGGAAACACCTTGACCTACGCCGCG 944

QY 2478 CGTGTCAAGGA 2488
    |||||
Db 943 CGGGCCAAGAA 933

RESULT 14
US-11-096-568A-4896
; Sequence 4896, Application US/11096568A
; Publication No. US20060048240A1
; GENERAL INFORMATION:
; APPLICANT: Alexandrov, Nikolai et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
```

```
; TITLE OF INVENTION: Therby
; FILE REFERENCE: 2750-1592PUS2
; CURRENT APPLICATION NUMBER: US/11/096,568A
; CURRENT FILING DATE: 2005-04-01
; NUMBER OF SEQ ID NOS: 34471
; SEQ ID NO 4896
; LENGTH: 1297
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1297)
; OTHER INFORMATION: Ceres Seq. ID no. 14304950
US-11-096-568A-4896
```

```
Query Match      1.8%; Score 54.6; DB 18; Length 1297;
Best Local Similarity 54.3%; Pred. No. 0.0017;
Matches 133; Conservative 0; Mismatches 109; Indels 3; Gaps 1;

QY 2258 TATGAGGGTCCGAGATTAAACAAATCGCTGCTGGCCCTCAAGGAGTGCATTCGTGCGTT 2317
    |||||
Db 46 TCTTGAGGGTGCATATATAACCGGTCTCTTCTTGCACTAAGCAGCTGTATTAAATTCCT 105

QY 2318 GGGCAACAGTCGGCCACTTGCCTTCCGTGCTCTCCAAATCACCAGGTGCTGCGCGA 2377
    |||||
Db 106 GGTAGAAGGCAAGAAACACATTCCTATACCGAAACTCAAACTCACTCAACTTCTCAAGGA 165

QY 2378 CTCGTTTCATTCGGAGAGAGCAAGACGTCGATGATAGCCATGATCTCCCGGAGCTTAG 2437
    |||||
Db 166 TTCACATA---GGAGGAACCTTGTAAACACTGTCTCATGTTGCAACATAAGCCCAAGTAACCT 222

QY 2438 CTCCTCGAGCAGCAGCTCAACACGCTGCGCTATGCGGATCGTGTCAAGGAGCTGCTGGT 2497
    |||||
Db 223 CTCATTTGTTGAACTCTCAGAACACAGTTCATTGGGCTGATAGCAAGAGATTTCGGGC 282

QY 2498 CAAGG 2502
    |||||
Db 283 AAAGG 287
```

```
RESULT 15
US-10-932-182A-84
; Sequence 84, Application US/10932182A
; Publication No. US20060046253A1
; GENERAL INFORMATION:
; APPLICANT: NAKAO, YOSHIHIRO
; APPLICANT: NAKAMURA, NORIHIISA
; APPLICANT: KODAWA, YUKIKO
; APPLICANT: FUJIMURA, TOMOKO
; APPLICANT: ASHIKARI, TOSHIHIKO
; TITLE OF INVENTION: METHODS FOR ANALYZING GENES OF INDUSTRIAL YEASTS
; FILE REFERENCE: 030685-043
; CURRENT APPLICATION NUMBER: US/10/932,182A
; CURRENT FILING DATE: 2004-09-02
; NUMBER OF SEQ ID NOS: 197023
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 84
; LENGTH: 3330
; TYPE: DNA
; ORGANISM: Saccharomyces pastorianus
US-10-932-182A-84
```

```
Query Match      1.7%; Score 52.4; DB 11; Length 3330;
Best Local Similarity 53.0%; Pred. No. 0.0074;
Matches 160; Conservative 0; Mismatches 136; Indels 6; Gaps 2;

QY 2184 GGCAGATTCTCGTTTCATCGATCTGCGGGCAATGAGCGGGCGTGGACACTTCTCTCGGCC 2243
    |||||
Db 934 GGTAAAGTTGAATTTGGTGGATTGGCTGGCAGTGA---GAACTTAACAGATCAGGTGCG 990

QY 2244 GATCGGAGAGCGGTATCGAGGGTCCAGAGATTAACAATCGCTGCTGCGCCCTCAAGGAG 2303
    |||||
Db 991 GAAACAAAAGAGCTCAAGAGCTGGACTGATTAACAATCACTGTTGACTTTAGTTCGT 1050
```





Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	454.6	14.8	1292	3	US-09-594-669-1	Sequence 1, Appli
2	454.6	14.8	1292	3	US-10-112-432-1	Sequence 1, Appli
3	454.6	14.8	1420	3	US-09-594-669-7	Sequence 7, Appli
4	454.6	14.8	1421	3	US-10-112-432-7	Sequence 7, Appli
5	454.6	14.8	2172	3	US-09-594-669-13	Sequence 13, Appli
6	454.6	14.8	2172	3	US-10-112-432-13	Sequence 13, Appli
7	449.8	14.6	1175	3	US-09-594-669-3	Sequence 3, Appli
8	449.8	14.6	1175	3	US-10-112-432-3	Sequence 3, Appli
9	449.8	14.6	1304	3	US-09-594-669-9	Sequence 9, Appli
10	449.8	14.6	1304	3	US-10-112-432-9	Sequence 9, Appli
11	449.8	14.6	2740	3	US-09-594-669-15	Sequence 15, Appli
12	449.8	14.6	2740	3	US-09-595-684B-32	Sequence 32, Appl
13	449.8	14.6	2740	3	US-10-112-432-15	Sequence 15, Appli
14	449.8	14.6	2740	3	US-09-849-602-5	Sequence 5, Appli
15	425.6	13.8	1088	3	US-09-594-669-5	Sequence 5, Appli
16	425.6	13.8	1088	3	US-10-112-432-5	Sequence 5, Appli
17	425.6	13.8	1217	3	US-09-594-669-11	Sequence 11, Appli
18	425.6	13.8	1217	3	US-10-112-432-11	Sequence 11, Appli
19	412.2	13.4	417	3	US-09-270-767-2183	Sequence 2183, Ap
20	412.2	13.4	417	3	US-09-270-767-17465	Sequence 17465, A
21	404.4	13.1	702	3	US-09-270-767-1072	Sequence 1072, A
22	404.4	13.1	1774	3	US-09-270-767-14823	Sequence 14823, A
23	396.6	12.9	2905	3	US-09-595-684B-24	Sequence 24, Appli
24	395.8	12.8	4617	3	US-09-620-312D-294	Sequence 294, App



Qy	2080	ATGCTGCCGCAACATCCGGCCAGACGTCGGCCAACTCCAATTCCTCGGTTTCGACACGCG	2139
Db	796	CGCGCTGCAGAACCTCTGGGCAGACATTTGCAACTCCAATTCCTCCGCTCCACACGCT	855
Qy	2140	TTTTCCAGATTGTCTGCGGCCGACGGCTCGAGAAAGATCCATGGCAAGTTCCTCGTTCA	2199
Db	856	GC TTCCAAATTA TTCTTCGAGCTAAAG-----GAGAA TGCATGGCAAGTTC TCTTTGG	909
Qy	2200	TCGATCTGGCGGGCAATGAGCGGGCGTGGA CACTTCTCGGCCGATCGGCAGACGCGTA	2259
Db	910	TAGATCTGGCAGGGAATGAGCGGGCGCGACACTTCCAGTGTCTGACCGGCAGACCCGCA	969
Qy	2260	TGGAGGGTGGCAGATTAA CAAATCGTCTGCTGGCCCTCAGAGAGTGATTCGTGCGTTGG	2319
Db	970	TGGAGGGCGCGAAATCAACAAGAGTCTCTTAGCCCTTGAAGGAGTGATCAGGGCCCCGTG	1029
Qy	2320	GCAACAGTCGGCCCACTTGCGCTTCCTCGTGTCTCCAACCTCACCCAGGTGCTCGCGACT	2379
Db	1030	GACAGACAGGCTCACACCCGTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGACT	1089
Qy	2380	CGTTCA TTGGCGAGAAGACGATGTCATGATAGCCATGATCTCCCGGGGACTTAGCT	2439
Db	1090	CTTTCA TTGGGAGAACTCTAGGACTTTGCATGATTGCCAGATCTCACAGGCATTAAGCT	1149
Qy	2440	CCTCGGAGCA CAGCTCAACAGCTGCGCTATGCGGATCGTGTCAAGGAGCTG	2492
Db	1150	CTGTGGAATATAC TT TTAACACCTCGAGATATG CAGACAGGGTCAAGGAGCTG	1202

### RESULT 3

US-09-594-669-7  
 ; Sequence 7, Application US/09594669  
 ; Patent No. 6331424  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Beraud, Christophe  
 ; APPLICANT: Sakowicz, Roman  
 ; TITLE OF INVENTION: No. 6331424el motor proteins and methods for  
 ; TITLE OF INVENTION: their use

Query Match	14.8%	Score 454.6	DB 3	Length 1420
Best Local Similarity	62.6%	Pred. No. 6.5e-107		
Matches 747	Conservative 0	Mismatches 434	Indels 12	Gaps 2
QY	1300	TGAAGAGGTGGAGCGACTCAAGGAGAAATCGCGAAGACGACGCCGCCGACAGCCGAGA	1359	
DB	21	TGAAGGAAGTGGAAAAAATGAGNACAAACGAGNAGAGAGAGAGGCCCCAGAACTCTGAAA	80	
QY	1360	TGAAGGAGGAGAGAGTGGCGCTGATGAACCAGAGATCGGGGAATCCAAACTGGGAGACGG	1419	
DB	81	TGAGAAATGAAGAGAG-----CTCAGAGATATGACAGTAGTTTTTCCAAACTGGGAATTTG	134	
QY	1420	CGCAAAATGAATACGCGAATATCAGAGCAGCTGGAAATTTGTGCGCTGCTCGATGGCCAGG	1479	
DB	135	CCCGAATGATTAAGAAATTTGGGCTACTTTGGAATCTCATCTACTATGACTATC	194	
QY	1480	CCGTGCGATGACCATCAGATCACAGTGTGGCTGGCGAAGCGTCCCATPAGCCGCAAGGAGG	1539	
DB	195	CTATCGAAGAGCACAGAATAATGTCTGTGTGTAGGAAACGCCCACTGAATTAAGCAAGAAAT	254	

RESIT.T 4

RESULT 4  
US-10-112-432-7  
; Sequence 7, Application US/10112432  
; Patent No. 6638754  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: No. 6638754el

```

; TITLE OF INVENTION: their use
; FILE REFERENCE: 1042A
; CURRENT APPLICATION NUMBER: US/10/112,432
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/724,215
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 09/314,464
; PRIOR FILING DATE: 1999-05-18
; PRIOR APPLICATION NUMBER: 09/594,669
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1421
; TYPE: DNA
; ORGANISM: Human
; US-10-112-432-7

Query Match      14.8%; Score 454.6; DB 3; Length 1421;
Best Local Similarity 62.6%; Pred. No. 6.5e-107;
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;

QY 1300 TGAAGAGGTGGAGCGACTGAAGGAGAAATCGCGAAGCGGACGCGCCCGACAGCCCGAGA 1359
DB 22 TGAAGGAAGTGGAAAAAATGAAGAACAAAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 81
QY 1360 TGAAGGAGGAGAGGTGGCGCTGATGAACACGAGGATCCGGGCAATCCAACTGGGAGACGG 1419
DB 82 TGAAGATGAAGAGAG-----CTCAGGAGATATGACAGTAGTATTTTCCAACTGGGAATTTG 135
QY 1420 CGCAATGATACGCGAATATCAGACGACGCTGGGAATTTGTGCGCGTCTCGATGGCCAGG 1479
DB 136 CCGAATGATTAAGAAATTTCCGGCTACTTTGGAAATGTCATCCACTTACTATGACTATC 195
QY 1480 CCGTCGATGACCATCAGATCAAGTGTGCGGCGCAAGCGTCCCATTAAGCGGCAAGGAGG 1539
DB 196 CTATCCAGAGCAGCAAGATATGTGCTGTGTAGGAAACGCCCACTGAATAAGCAAGAT 255
QY 1540 TCAATCGCAGGAGATCGATGTCATTTTGGTCCGCGCAAGGACATGCTCATCGTGACG 1599
DB 256 TGCCCAAGAAAGAAATTTGATGATTTCCATTCCTTAGCAAGTGTCTCTCTTGTGTACATG 315
QY 1600 AGCCGCGCAGCAAGGTGCGACCTCACCAAGTTCCTTGGAAGAACCAAGTTTCGTTCGACT 1659
DB 316 AACCCAAGTTGAAGTGGACTTAAACAAGATATCGGAGAACCAAGATTCGTCTTTGACT 375
QY 1660 AGCCCTTCAACGACAGCTGCGCAATGCCATGGTATACAAATACACAGCAAGCCGTTGG 1719
DB 376 TTGCATTTGATGAACAGCTTCGAATGAAGTTGTCTACAGGTTTCACAGCAAGGCCACTGG 435
QY 1720 TGAAGAACATTTTCGAGGGCGGAATGGCGAGCTGCTTCGCTACGGCCAGACGGGATCGG 1779
DB 436 TACAGACAATCTTTGAAGGTGGAAAGCAACTTGTTTTGCATATGGCCGACAGAGGAAGTG 495
QY 1780 GCAAAAACGACACCACTGGCGGTGAGTTTAAATGAAAGGTGTCAGGACTGCAAGAACGGCA 1839
DB 496 GCAAGACACATACTATGCGGCGAGACCTCTCTGGAAAGCCAGAAATGCATCCAAAGGGA 555
QY 1840 TCTAGCCCATGCGGCGCAAGGATGTCTTTGTGACCCCTGAATPATCCCGCGTTACCGGCCA 1899
DB 556 TCTATGCCATGCGCTCCCGGAGCGTCTTCTCTCTGMAAGAAATCAACCTCTACCGGAAGT 615
QY 1900 TGAATCTAGTCTCTCGGCCAGTTTCTTTGAGATTTACAGTGGCAAGGTCTTGCATCTTC 1959
DB 616 TGGGCTCGGAAGTCTATGTGACATCTTTCGAGATCTACAATGGGAAGCTGTGTTGACCTGC 675
QY 1960 TGTCCGACAAAGCAGAAACTGCGCGTCTCGGAGGATGGTAAACAGCAAGATGCAAGTGTGG 2019
DB 676 TCAACAAGAGGCCAAGCTGCGCGTCTGGAGGATGGCAGCAACAGGTGCAAGTGTGGTGG 735
QY 2020 GACTCACCGAAGAGTGTGATGATGCGGTCTGAGGAGGTACTGAAAGCTCATTCAGCACGGCA 2079

```

```
Db      752 CTATCGAGACCAAGATATGTGTGTGTAGAAACGCCCTGAAATAAGCAAGAT 811
Qy      1540 TCAATCGAAGGAGATCGATGTCATTTTGGTGGCGCGCAAGGACATGCTCATCTGTGCAG 1599
Db      812 TGGCCAAAGAAAGAAATGATGTGATTTCCATTCTCTAGCAAGTGTCTCTTGTGTACATG 871
Qy      1600 AGCGCGCAGCAAGGTGACCTCACCAAGTTCCTGGAGAACCAAGTTTGGCTTCGACT 1659
Db      872 AACCCAAATGTAAGAGTGACTTAACAAGATATCTGGAGAACCAAGCATTTCTGCTTGA 931
Qy      1660 ACGCCTTCAACGACACGCTGCACAAATCCATGTTATACAAATACACAGCCAAAGCGTTGG 1719
Db      932 TTGCAATTGATGAACAGCTTCGAATGAGTTGTCTACAGTTTACAGCAAGGCCACTGG 991
Qy      1720 TGAACAACATTTTCGAGGGCGGAATGCGAGCTGCTTCGCTACGGCCAGACGGGATCGG 1779
Db      992 TACAGACAAATCTTTGAAGGTGGAAGAACAACTTGTGTCATATGGCCAGCAGGAAGTG 1051
Qy      1780 GCAAAACGACACACCATCGGCGGTGAGTTTAATGGAAGGTGCAGGACTGCAGAACGGCA 1839
Db      1052 GCAAGACACATACTATGCGCGGAGACCTCTCTGGGAAAGGCCAGAAATGCAATGCCAAAGGGA 1111
Qy      1840 TCTACGCGCATGGCGCCAGGATGCTTTGTGACCTCGCTGAATATATCGCGGTACCGGCCCA 1899
Db      1112 TCTATGCCATGGCCTCCCGGGACGCTTCTCTCTGAAGAAATCAACCTGCTACCGGAAGT 1171
Qy      1900 TGAATCTAGTCTCGGCCAGTTCCTTTTGAGATTTTACAGTGGCAAGGCTTTCGATCTTTC 1959
Db      1172 TGGGCTTGGAAATCTATGTGACATTTCTCGAGATCTACAAATGGAGCTGTTTGAACCTGC 1231
Qy      1960 TGTCCGACAGCAGAACTCGGCGTCTCGAGAGTGTAAACAGCAAGTGCAGGTGGTG 2019
Db      1232 TCAACAGAAAGGCCAAGCTCGCGGTGCTGGAGGATGGCAAGCAAGGTGCAAGTGGTG 1291
Qy      2020 GACTCACCGAGAGGTGCTCGATGGCGTCGAGAGGTACTGAACTCATCCAGCAGCGCA 2079
Db      1292 GGCTGCAGGAGCATCTGTTAACTCTCTGATGATGTCATCAAGATGATCGACATGGGCA 1351
Qy      2080 ATGCTGCCGGAACATCCGGCCAGACGCTCGGCCAACTCCAATTCGTGCGCTTCGACGCGG 2139
Db      1352 GCGCCTGCAGAACCTCTGGGCAGACATTTTGCCTCACTCCAAATTCCTCCGCTCCACGCGT 1411
Qy      2140 TTTTCCAGATGTGCTCGCGCGCAGGCTCGAGAGATCCATGGAAGTTCCTGTTCA 2199
Db      1412 GCTTCCAAATTAATCTTCGAGCTAAAGG-----GAGAATGTCATGGCAAGTCTCTTTGG 1465
Qy      2200 TCGATCTGGCGGCAATAGCGGGGCTGGACACTTCTCGCGGATCGGCAGACGGTA 2259
Db      1466 TAGATCTGGCAGGGAATGAGCGGCGGACACTTCCAGTGTGACCGGCAGACCGCA 1525
Qy      2260 TGGAGGTGCGGAGATTAACAATTCGCTGGCCCTCAAGGAGTGCATTCGTGCGTGG 2319
Db      1526 TGGAGGCGCGAGAAATCAACAAGATCTCTTAGCCCTGAGGAGTGCATCAGGGCCCTGG 1585
Qy      2320 GCAAAACGTCCGCCCATCTTCCCTTCGTGTCTCCAACTCAACCAAGGTGCTCGCGACT 2379
Db      1586 GACAGAAACAGGCTCAACCCCGTTCGTGAGAGCAAGCTGACACAGGTGCTGAGGACT 1645
Qy      2380 CGTTTCAATGCGAGAGAGCAAGACGTGCATGATGACCATGATCTCGCCGGACTTAGCT 2439
Db      1646 CTTTCATTTGGGAGAACTCTAGGACTTGCATGATGTCACAGATCTCACAGGCGATAGCT 1705
Qy      2440 CTTGCGAGCACCGCTCAACACGCTGCTATCGGATCGTGTCAAGGAGCTG 2492
Db      1706 CTTGTGATATATCTTTAAACACCTTGATATGCAGACAGGGTCAAGGAGCTG 1758
```

RESULT 6  
US-10-112-432-13  
; Sequence 13, Application US/10112432  
; Patent No. 6638754  
; GENERAL INFORMATION:

```
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: No. 6638754el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1042A  
; CURRENT APPLICATION NUMBER: US/10/112,432  
; CURRENT FILING DATE: 2002-03-29  
; PRIOR APPLICATION NUMBER: 09/724,215  
; PRIOR FILING DATE: 2000-11-28  
; PRIOR APPLICATION NUMBER: 09/295,612  
; PRIOR FILING DATE: 1999-04-20  
; PRIOR APPLICATION NUMBER: 09/314,464  
; PRIOR FILING DATE: 1999-05-18  
; PRIOR APPLICATION NUMBER: 09/594,669  
; PRIOR FILING DATE: 2000-06-15  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 2172  
; TYPE: DNA  
; ORGANISM: Human  
; US-10-112-432-13
```

```
Query Match      14.8%; Score 454.6; DB 3; Length 2172;  
Best Local Similarity 62.8%; Pred. No. 8e-107;  
Matches 747; Conservative 0; Mismatches 434; Indels 12; Gaps 2;  
  
Qy      1300 TGAAGAGGTGGAGCGACATGACGAGAAATCGCGAGAAAGCGACGCGCCGACAGCCGAGA 1359  
Db      578 TGAAGAGGTGGAAAGAAATGAAAGAACAGCGAGAGAGAAAGAGCCGAACTCTGAAA 637  
Qy      1360 TGAAGAGAGAAAGGTGGCGCTGATGAACAGGATCCGGGCAATCCAAACTCGGAGACGG 1419  
Db      638 TGAGAAATGAAGAGAG-----CTCAGGAGTATCACAGTAGTTTTTCCAAACTGGGAATTTG 691  
Qy      1420 CGCAATATGATACCGGAATATCAGAGCACGCTGGAATTTGTGCGCTGCTCGATGGCCAGG 1479  
Db      692 CCGCAATGAAATGAAGAAATTTTCGGGCTACTTTGGAATGTCTATCCACTTACTATGACTGATC 751  
Qy      1480 CCGTCCATGACCATCAGATCAGATCAGATGCTGCGTGGCAAGGTGCCATTTAGCCGCAAGGAGG 1539  
Db      752 CTATCGAAGAGCACAGAAATATGTGTCTGTGTAGGAAAGCCCACTGAATAAGCAAGAT 811  
Qy      1540 TCAATCGAAGGAGATCGATGTCATTTTCGGTGGCGCGCAAGGACATGCTCATCTGTGCAGC 1599  
Db      812 TGGCCAAAGAAAGAAATTTGATGTGATTTCCATTCTTAGCAAGTGTCTCTCTTGTGTACATG 871  
Qy      1600 AGCGCGCAGCAGAGGTGCGACTCACCAAGTTCTCTGGAGAACCAAGTTTTCGCTTCGACT 1659  
Db      872 AAGCCCAAGTTGAAGTGGGACTTTAAACAAAGTATCTGGAGAACCAAGCATTTCTGCTTTGACT 931  
Qy      1660 ACGCCTTCAACGACACGCTGCGACAAATGCCATGTTATACAAATACACAGCCAAAGCGTTGG 1719  
Db      932 TTGCAATTGATGAACAGCTTCGAATGAAAGTTGTCTACAGTTTACAGCAAGGCCACTGG 991  
Qy      1720 TGAACAACATTTTCGAGGGCGGAATGGCGACGTGCTTCGCTACCGCCAGACGGGATCGG 1779  
Db      992 TACAGACAAATCTTTGAAGGTGGAAGAACAACTTGTGTCATATGGCCAGCAGGAAGTG 1051  
Qy      1780 GCAAAACGACACCATCGGCGGTGAGTTTAATGGAAGGTGCAGGACTGCAGAACAGCGCA 1839  
Db      1052 GCAAGACATATCTATGGCGGAGACCTCTCTGGGAAAGGCCAGAAATGCAATGCCAAAGGGA 1111  
Qy      1840 TCTAGCCCATGGCGCCAGGATGCTTTGTGACCTCGCTGAATATATCGCGGTACCGGCCCA 1899  
Db      1112 TCTATGCCATGGCCTCCCGGGACGCTTCTCTCTGAAGAAATCAACCTGCTACCGGAAGT 1171  
Qy      1900 TGAATCTAGTCTCGGCCAGTTCCTTTTGAGATTTTACAGTGGCAAGGCTTTCGATCTTTC 1959  
Db      1172 TGGGCTTGGAAATCTATGTGACATTTCTCGAGATCTACAAATGGAGCTGTTTGAACCTGC 1231  
Qy      1960 TGTCCGACAGCAGAACTCGGCGTCTCGAGAGTGTAAACAGCAAGTGCAGGTGGTG 2019
```

Db 1232 TCAACAAGAGGCAAGCTGCGCTGCTGGAGATGGCAAGCAACAGAGTGCAAGTGGTGG 1291  
Qy 2020 GACTCACAGAGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2079  
Db 1292 GCGTCAGAGAGATCTGGTTAACTCTGCTGATGATGATGATGATGATGATGATGATGATG 1351  
Qy 2080 ATGCTGCCCGAATCTCGGCGCAGAGAGTGGGCAATCTCCAAATTCGTCGGTTCGACGCGG 2139  
Db 1352 GCGCTGCAGAACCTCTGGGCGAGACATTTGGCCAACTCCAAATTCCTCCGCTCCAGCGGT 1411  
Qy 2140 TTTTTCAGATGTTGCTGGGCGCGAGAGTGGGCAATCTCCAAATTCGTCGGTTCGACGCGT 2199  
Db 1412 GCTTCCAAATTAATTTCTGAGCTAAAG-----GAGAATGATGCGCAAGTTCTCTTGG 1465  
Qy 2200 TCGATCTGCGGCGCAATGAGCGGCGTGGACACTCTCTCGCCCATCGGCGAGACGCTA 2259  
Db 1466 TAGATCTGCGAGGGAATGAGCGGCGGGAACACTTTCAGTCTGACCGGCGAGACCCGCA 1525  
Qy 2260 TGGAGGTGCGGAGATTAACAAATCGCTGCTGGCCCTCAAGGAGTGCAATTCGTGGGTGG 2319  
Db 1526 TGGAGGCGCGAGAAATCAACAGAGTCTTTAGCCCTGAGGAGTGATCAGGCGCTGG 1585  
Qy 2320 GCAACAGTGGCGCACTTGGTCCCTTCGCTGCTCAAACTCACAGAGTGGTGGCGACT 2379  
Db 1586 GACAGAAACAAGGCTCACACCCGTTCCGCTGAGAGCAAGCTGACACAGGTGCTGAGGACT 1645  
Qy 2380 CGTTCAATGGCGAGAGCAAGAGTGCATGATAGCCATGATCTCGCGGCACTTAGCT 2439  
Db 1646 CTTTCAATGGGAGAACTCTAGGACTTGCATGATTTGCCACGATCTCACCGGCATAAGCT 1705  
Qy 2440 CTTGCGAGCACAGCTCAACAGCTGCGCTATGCGGATGCTCAAGGAGCTG 2492  
Db 1706 CTTGTGAATATCTTTAAACCCCTGAGATATGACAGAGGTCAAGGAGCTG 1758

## RESULT 7

US-09-594-669-3  
; Sequence 3, Application US/09594669  
; Patent No. 6331424  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: No. 6331424el motor proteins and methods for  
; TITLE OF INVENTION: their use  
; FILE REFERENCE: 1042  
; CURRENT APPLICATION NUMBER: US/09/594,669  
; CURRENT FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 09/295,612  
; PRIOR FILING DATE: 1999-04-20  
; PRIOR APPLICATION NUMBER: US 09/314,464  
; PRIOR FILING DATE: 1999-05-18  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1175  
; TYPE: DNA  
; ORGANISM: Human  
US-09-594-669-3

Query Match 14.6%; Score 449.8; DB 3; Length 1175;  
Best Local Similarity 64.0%; Pred. No. 1e-105;  
Matches 698; Conservative 0; Mismatches 387; Indels 6; Gaps 1;

Qy 1402 ATCCAACTGGAGAGCGGCAATATACCGGAATATCAGAGACGCTGGAATTTGTGC 1461  
Db 1 ATGCAAACTGGGAATTTGCCGGAATGATTAAGGAATTTCCGGCTACTTTGGAATGTATC 60  
Qy 1462 CGCTCTCGATGGCGGCGGCTGATGACCATCAGATCAGATGCTGGCGCAAGCTC 1521  
Db 61 CACTTACTATGACTGATCTTATCGAGAGACAGATATGCTGTGTGTAGGAACGCC 120  
Qy 1522 CCAATPAGCCGCAAGGAGTCAATCGAAGGAGATCGATGTCAATTTCCGTGCGCGCAAGG 1581

Db 121 CACTGAATAGCAAGAATTGGCCAAGAAAGAAATTTGATGTGATTTTCAATTCCTAGCAAGT 180  
Qy 1582 ACATGCTCATCGTGCACGAGCGCGCAGCAAGGTGCGACTCACCAGTTTCTCTGAGAAACC 1641  
Db 181 GTCTCTCTTGTGATGAAACCAAGTTGAAAGTGGACTTAAACAAGTATCTGGAGAAC 240  
Qy 1642 ACAAGTTTTCGCTTCGACTTACGCTTCAACGACACGTCGCAAAATGCCATGTATACAAAT 1701  
Db 241 AAGCATTTCTGTTGACTTTGCTGATGAAACAGCTTCGAAATGAAGTTGCTTACAGGT 300  
Qy 1702 ACACAGCCCAAGCCGTTGCTGAAACCAATTTTCGAGGCGGGAATGGCGAGTCTTCGCT 1761  
Db 301 TCACAGCAAGGCCACTGGTACAGCAATCTTTTGAAGGTGGAAAGCAACTTTGTTTGCAT 360  
Qy 1762 ACAGCCAGAGCGGATCGGCGAAACGACACCATTCGCGGCTGAGTTTAAATGGAAAGGTGC 1821  
Db 361 ATGSCCAGACAGGAAGTGGCAAGACACATATATGGGCGGAGACTCTCTGGGAAAGCCC 420  
Qy 1822 AGGACTGCAAGAACCGGATCTACGCCATGGCGGCAAGGATGTCTTTGTGACCCCTGAATA 1881  
Db 421 AAGATGCAATCAAAAGGAGATCTATGCAATGGCTCCCGGAGCTCTTCTCTGAGAAATC 480  
Qy 1882 TGCCGCGTTACCGCGCCATGAATCTAGTCTGCTCGGCGAGTTTCTTTGAGATTTACAGTG 1941  
Db 481 AACCTGCTACCGGAAGTTGGGCTTGGAGTCTATGTGACATTTCTTCGAGATCTACAATG 540  
Qy 1942 GCAAGGTCTTCGATCTTCTGTCGCAAGCAAGAAACTGCGGCTCTCTGAGGATGGTAAAC 2001  
Db 541 GGAAGCTGTTTGAACCTGCTCAACAAGAAAGGCAAGCTGCGCTGCTGGAGGATGGCAAGC 600  
Qy 2002 AGCAAGTGCAGGTGGTGGGACTCACCGAGAGGTGGTTCGATGGGCTCGAGGAGTACTGA 2061  
Db 601 AACAGGTGCAAGTGGTGGGCTCGAGGAGCATCTGGTTAACTCTGCTGATGATGTCAATCA 660  
Qy 2062 AGCTCATCCAGCACGCGCAATGCTGCCCGAAACATCCGGCGCAGACGCTCGGCCCAACTCCAATT 2121  
Db 661 AGATGATCGACATGGCGAGCGCTGCAGAACCTCTGGGCGAGACATTTGCCAACTCCAATT 720  
Qy 2122 CGTCGCTTGCAGCGCGTTCCTCAGATTTGCTGCGGCGCAGCGGCTCGACGAGATCC 2181  
Db 721 CTTCCGCTCCACGCGTGTCTCCAAATTAATTTCTCGAGCTTAAAG-----GAGAAATGC 774  
Qy 2182 ATGCAAGTTCTCTTCTCATCGATCTGCGGCGCAATGAGCGGCGTGGACACTTCTCTCGG 2241  
Db 775 ATGCAAGTTCTCTTCTGATGATCTGCGAGGGAATGAGCGGCGCGGACACTTCCAGTG 834  
Qy 2242 CCGATCGGCGAGACGCTATGAGGGTCCCGAGATTAACAATCGCTCTGCGCCCTCAAGG 2301  
Db 835 CTGACCGCGAGACCGCATGGAGGCGCAGAAATCAACAAGAGTCTCTTAGCCCTGAAGG 894  
Qy 2302 AGTCATTTGCTGGTGGGCAACAGTCGGCCCACTTGGCCCTTCCGTGCTTCCAAACTCA 2361  
Db 895 AGTCATCAGGGCCCTGGGACAGAAACAAGGCTCACACCCGCTTCCGTGAGAGCAAGCTGA 954  
Qy 2362 CCAGGTGCTGCGGACTCTGTTTCAATGGCGAGAGAGCAAGACGTCATGATGCCATGA 2421  
Db 955 CACAGGTGCTGAGGACTCTTCTTATGGGAGAACTCTTAGACTTTGATGATGCCAGA 1014  
Qy 2422 TCTCGCGGACTTTAGTCTCTCGAGACACAGCTCAACAGCTGCGCTATCGGATCGTG 2481  
Db 1015 TCTCAACAGGACATAAGCTCTCTGTGAATATACTTTTAAACACCCCTGAGATATGCACAGGG 1074  
Qy 2482 TCAAGGAGCTG 2492  
Db 1075 TCAAGGAGCTG 1085

## RESULT 8

US-10-112-432-3  
; Sequence 3, Application US/10112432  
; Patent No. 6638754  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe



APPLICANT: Sakowicz, Roman  
TITLE OF INVENTION: No. 631424el motor proteins and methods for  
TITLE OF INVENTION: their use

FILE REFERENCE: 1042A

CURRENT APPLICATION NUMBER: US/10/112.432

PRIOR FILING DATE: 2002-03-29

PRIOR APPLICATION NUMBER: 09/724,215

PRIOR FILING DATE: 2000-11-28

PRIOR APPLICATION NUMBER: 09/295,612

PRIOR FILING DATE: 1999-04-20

PRIOR APPLICATION NUMBER: 09/314,464

PRIOR FILING DATE: 1999-05-18

PRIOR APPLICATION NUMBER: 09/594,669

PRIOR FILING DATE: 2000-06-15

NUMBER OF SEQ ID NOS: 16

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 3

LENGTH: 1175

TYPE: DNA

ORGANISM: Human

US-10-112-432-3

Query Match 14.6%; Score 449.8; DB 3; Length 1175;

Best Local Similarity 64.0%; Pred. No. 1e-105;

Matches 698; Conservative 0; Mismatches 387; Indels 6; Gaps 1;

```
QY 1402 ATCAAACTGGGAGCGGCAAAATGATACGCGAATATACAGACACGCTGGAATTTGTGC 1461
Db 1 ATGCAAACTGGGAATTTGCCCGAATGATTAAGAAATTTTCGGGTACTTTGGAATGTCATC 60

QY 1462 CGCTGCTCGATGCCAGGCGTCGATGACCATCAGATCAGATGCTGCGTGGCGCAAGCTC 1521
Db 61 CACTTACTATGATCTGCTATCGAAGAGCACAGATATATGCTGTGTGTAGGAACGCC 120

QY 1522 CCATTAGCCCAAGGAGGTCAATCGCAAGGAGATCGATGTCATTTTCGGTCCCGCAAGG 1581
Db 121 CACTGAATAGCAAGAAATGGCCCAAGAAATGATGATTTCCATTTCTAGCAAGT 180

QY 1582 ACATGCTCATCTGTCAGAGCGCGGAGCAAGGTGACCTCACAAATTCCTGGGAACC 1641
Db 181 GTCTCTCTTTGTACATGAACCAAGTTGAAAGTGAGCTTAAACAAAGTATCTGGGAACC 240

QY 1642 ACAAGTTTCGCTCGACTAGCCCTTCAACGACAGTGGCAGATGCGATGATGATACAAAT 1701
Db 241 AAGCATCTGCTTTGATCTTTCATTTGATGAACACAGCTTCGAATGAAGTTGTCTACAGT 300

QY 1702 ACACAGCCCAAGCGTTTGGTGAAACCAATTTTCAGGCGCGAATGGCGAGCTGCTCGCCT 1761
Db 301 TCACAGCAAGGCACCTGGTACAGACATCTTTGAGGTGGAAAGCAACTGTTTGTGCAT 360

QY 1762 ACGGCGCAGAGGATCGGGCAAAACGACACATGCGGCGGTGAGTTTAAATGGAAAGGTGC 1821
Db 361 ATGGCCAGACAGGAAGTGGCAAGACACATACTATGGCGGAGACCTCTCTGGGAAGCCC 420

QY 1822 AGGACTGCAAGACGCGATCTAGCCATGGCGCCCAAGATGCTTTGTGACCTGAATA 1881
Db 421 AGAATGCATCAAGGAGTCTATGCGCATGGCCCTCCCGGAGCGCTTCTCTCTGAAGATTC 480

QY 1882 TGCCGCGTTACCGCGCCCAATGATCTAGTCTCTCGGCCAGTTTCTTTGAGATTTACAGT 1941
Db 481 AACCCTGCTACCGAAGTTGGGCTGGAAGTCTATGTGACATCTTCGAGATCTACATG 540

QY 1942 GCAAGTCTTCGATCTTCTGTCGCAAGCAGAAACTGGCGGCTCTGGAGGATGGTAAC 2001
Db 541 GGAAGCTGTTGACCTGCTCAACAAAGAGCCCAAGCTGCGGCTGGAGGATGGCAAGC 600

QY 2002 AGCAAGTGCAGGTGGGAGTCAACCGAGAGGTGTCATGCGCTGCGAGGAGTACTGA 2061
Db 601 AACAGTGCAGGTGGGAGTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCA 660

QY 2062 AGCTCATCCAGCAGCAATGCTGCGCGCAACATCCGCGCAGAGCTGCGCAACTCCAATT 2121
Db 661 AGATGATCAGATGGGAGGCGCTGCAGAACTCTTGGGAGACATTTGGCAACTCCAATT 720
```

```
QY 2122 CGTGGCTTGGACAGCCGTTTTCAGATTTGCTGCGGCGCAGGGCTCGACGAAGATCC 2181
Db 721 CTTCCGCTCCACACGCTGCTTCCAAATATTCTTCGAGCTAAAGG-----GAGAATGC 774

QY 2182 ATGGCAAGTCTCGTTTCATCGATCTGCGGCGCAATGAGCGGCGGTGGACACTTCTCTCG 2241
Db 775 ATGGCAAGTCTCTTTTGGTAGATCTGGCAGGGAATGAGCGAGCGCGGACACTTCCAGTG 834

QY 2242 CCGATCGGAGAGCGCTATGAGGCGTCCGAGATTAACAAATCGCTGCGGCCCTCAAGG 2301
Db 835 CTGACCGGAGAGCCCGCATGAGGCGCAGAAATCAACAGAGTCTCTTAGCCCTGAGG 894

QY 2302 AGTGCAATTCGCTGTTGGGCAACAGTCCGCCACCTTTCCTGCTTCCAAACTCA 2361
Db 895 AGTGCAATCAGGCGCTTGGGACAGAACAGGCTCACACCCCGTTCCTGAGAGCAAGCTGA 954

QY 2362 CCGAGTGTGCGGAGCTGTTTCATTTGGCGAGAGAGCAAGCTGTCATGATAGCCATGA 2421
Db 955 CACAGTGTGAGGAGCTCTCTTCAATTTGGGAGAACTCTTAGGACTTTCATGATTTGCCACGA 1014

QY 2422 TCTCGCGGAGCTTAGCTCTCTGCGAGCACACGCTCAACAGCTGCGCTATGCGGATCGTG 2481
Db 1015 TCTCACCAGGCATAGCTCTCTGTAATATATCTTTAAACACCTTGAGATATGAGACAGG 1074

QY 2482 TCAAGGAGCTG 2492
Db 1075 TCAAGGAGCTG 1085
```

#### RESULT 9

US-09-594-669-9

Sequence 9, Application US/09594669

Patent No. 6331424

GENERAL INFORMATION:

APPLICANT: Beraud, Christophe

APPLICANT: Sakowicz, Roman

TITLE OF INVENTION: No. 6331424el motor proteins and methods for

TITLE OF INVENTION: their use

FILE REFERENCE: 1042

CURRENT APPLICATION NUMBER: US/09/594,669

PRIOR FILING DATE: 2000-06-15

PRIOR APPLICATION NUMBER: US 09/295,612

PRIOR FILING DATE: 1999-04-20

PRIOR APPLICATION NUMBER: US 09/314,464

PRIOR FILING DATE: 1999-05-18

NUMBER OF SEQ ID NOS: 16

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 9

LENGTH: 1304

TYPE: DNA

ORGANISM: Human

US-09-594-669-9

Query Match 14.6%; Score 449.8; DB 3; Length 1304;  
Best Local Similarity 64.0%; Pred. No. 1.1e-105;  
Matches 698; Conservative 0; Mismatches 387; Indels 6; Gaps 1;

```
QY 1402 ATCAAACTGGGAGCGGCAAAATGATACGCGAATATACAGACACGCTGGAATTTGTGC 1461
Db 1 ATGCAAACTGGGAATTTGCCCGAATGATTAAGAAATTTTCGGGCTACTTTGGAATGTCATC 60

QY 1462 CGCTGCTCGATGCCAGGCGCTGATGACCATCAGATCAGATGCTGCGTGGCGCAAGCTC 1521
Db 61 CACTTACTATGATCTGATCTCTATCGAAGAGCACAGATATGCTGTGTGTAGGAACGCC 120

QY 1522 CCATTAGCGCGAAGGAGTCAATCCAGGAGATCGATGTCATTTTCGGTGGCGCAAGG 1581
Db 121 CACTGAATAGCAAGAAATGGCCCAAGAAATGATGATGATTTCCATTTCTTAGCAAGT 180

QY 1582 ACATGCTCATCTGTCAGAGCGCGCAGCAGCATGCTCACCAGTTCCTGGAGAAC 1641
Db 181 GTCTCTCTTGTGATCATGAACCAAGTTGAAAGTGAAGTGAAGTGAAGTATCTGGAGAAC 240
```

```
QY 1642 ACAAGTTTCGCTTCAAGTACGCTTCAACGACACGTCGACAAATGCCATGGTATACAAAT 1701
Db 241 AAGCATTTCTGCTTTGACTTTGATTTGATGAACACAGCTTCGAATGAAGTTGTCTACAGT 300
QY 1702 ACACAGCCAGCGGTTGGTGAACAAATTTTCGAGGGCGGAATGCGAGCTCTTCGCT 1761
Db 301 TCACAGCAGGCGCACTGGTACAGACAAATCTTTGAAGTTGGAAGAACAACTCTTTTGGCAT 360
QY 1762 ACGGCCAGAGGATCGGCGCAACACGACACCATGGCGGTGAGTTTAATGGAAGGTCG 1821
Db 361 ATGGCCAGACAGGAATGCGAGACACATACATATGGCGGAGACCTCTCTGGGAAGGCC 420
QY 1822 AGGACTCGAAGAACGGCACTTACGCCATGGCGGCCAAGGATGCTCTTTGTGACCCCTGAATA 1881
Db 421 AGAATGCATCCAAAGGATCTATGCATGGCCCTCCGGGAGCGTCTTCTCTGAAGATC 480
QY 1882 TGCCCGGTTACCGGCCCATGAATCTAGTGTCTCGGCCAGTTCTTTTGAGATTTACAGTG 1941
Db 481 AACCTGTCTACCGGAAGTTGGGCTTGAAGTCTATGTGACATTTCTTCGAGATCTACAATG 540
QY 1942 GCAAGCTCTTCGATCTTCTGTCGACAGCAAGCAAACTGGCGTCTGGAGATGTTAAAC 2001
Db 541 GGAAGCTGTTGACCTGTCTCAACAAAGAGGCCAAGCTGGCGTCTGGAGGATGGCAAGC 600
QY 2002 AGCAAGTGCAGGTGGTGGGACTCACCGAGAAAGGTGGTGCATGGCGTCTGAGGAGTACTGA 2061
Db 601 AACAGTGCAGTGGTGGGCTGACGAGCATCTGTTAACTCTGCTGATGATGTCATCA 660
QY 2062 AGCTCATCAGCAGCGCAATGTGCGCCGAAACATCGGCAGACGTCGCGCCAACTCCAAT 2121
Db 661 AGATGATCGACATGGCGAGCGCTGCAGAACCTCTGGCGAGACATTTGCCAACTCCAAT 720
QY 2122 CGTCGCTTCGACCGCTTTTTCAGATTTGCTCGCGCGGAGGCTCGAGAGATCC 2181
Db 721 CCTCCGCTCCCAAGCTGCTTCCAAATTAATCTTTCGAGCTAAAGG-----GAGAATGC 774
QY 2182 ATGGCAAGTCTCTGTTTCATCGATCTGGCGGCAATGAGCGGGGTGGACACTTCTCTCGG 2241
Db 775 ATGGCAAGTCTCTTGTGTAGATCTGGCAGGGAATGAGCGGCGCGACACTTCCAGTG 834
QY 2242 CCGATCGCAGACGCGTATGAGGGTGGCGAGATTAACAAATCGCTGCTGCCCTCAAGG 2301
Db 835 CTGACCGGCGAGACCGCATGGAGGCGCAGAAATCAACAAGATCTCTTAGCCCTGAGG 894
QY 2302 AGTCATTCGTGCGTGGGCAACAGTCGCGCCACTTCCCTTCCGTGTCTCCAAACTCA 2361
Db 895 AGTGATCAGGGCCCTGGGACAGAACAAAGGCTCACACCCGCTTCGCTGAGAGCAAGCTGA 954
QY 2362 CCCAGGTGCTCGCGACTCGTTTCATTGGCGGAGAGAGCAAGCTGCAATGATAGCATGA 2421
Db 955 CACAGGTGCTGAGGACTTCTTTCATTGGGGAGAACTTAGGACTTGCATGATTTGCCAGA 1014
QY 2422 TCTCGCGGACTTAGCTCTCGAGCAGACAGCTCAACACGCTGCGCTATCGGATCGTG 2481
Db 1015 TCTCACCAGGCATAAGCTCTCTGTAATATACTTTTAAACACCCCTGAGATATCAGACAGG 1074
QY 2482 TCAAGGAGCTG 2492
Db 1075 TCAAGGAGCTG 1085
```

## RESULT 10

```
US-10-112-432-9
; Sequence 9, Application US/10112432
; Patent No. 6638754
; GENERAL INFORMATION:
; APPLICANT: Berard, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: No. 6638754e1 motor proteins and methods for
; FILE OF INVENTION: their use
; FILE REFERENCE: 1042A
; CURRENT APPLICATION NUMBER: US/10/112.432
```

```
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/724,215
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 09/314,464
; PRIOR FILING DATE: 1999-05-18
; PRIOR APPLICATION NUMBER: 09/594,669
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 1304
; TYPE: DNA
; ORGANISM: Human
US-10-112-432-9
```

```
Query Match 14.6%; Score 449.8; DB 3; Length 1304;
Best Local Similarity 64.0%; Pred. No. 1.1e-105;
Matches 598; Conservative 0; Mismatches 387; Indels 6; Gaps 1;

QY 1402 ATCCAACTGGGAGCGCGCAATATACCGAATATACGAGCAGCTGGAATTTGTGC 1461
Db 1 ATGCAAACTGGGAATTTGCCCAATGATTAAAGAAATTTCCGGCTACTTTTGGAAATGTCATC 60
QY 1462 CGCTGCTCGATGGCCGCGCTCGATGACCATCAGATCAGATGCGTGGCGCAGCGTC 1521
Db 61 CACTTACTATGACTGATCTCTATCGAAGAGCAGACAAATATGTGTCTGTGTTAGGAAACGCC 120
QY 1522 CCATTAGCCGCAAGAGGTCAATCGCAAGGAGATCGATGTCAATTTCCGGTCGCGCAAGG 1581
Db 121 CACTGAATACGCAAGAAATTTGCCAAGAAATTTGATGATTTTCCATTTCTAGCAAGT 180
QY 1582 ACATGCTCATCGTGCACGAGCGCGCAGCAAGTTCGACCTCACCAGTTCCTGGGAACC 1641
Db 181 GTCTCCTCTTGGTACATGAAACCAAGTTGAAGTGGACTTAACAAGATATCTGGGAACC 240
QY 1642 ACAAGTTTCGCTTCGACTACGCTTCAACGACAGCTGGGCAATGCCATGGTATACAAAT 1701
Db 241 AAGCATTTCTGCTTTGACTTTGATGATGAACAGCTTTCGAATGAAGTTGTCTACAGGT 300
QY 1702 ACACAGCCAGCGCTTGGTGAACCAATTTTCGAGGGCGGAATGGCGAGCTTTCGCT 1761
Db 301 TCACAGCAAGGCCACTGGTACAGACAACTTTGAAGTGGAAAGCAACTTTGTTTGGAT 360
QY 1762 ACGGCCAGACGGGATCGGCGCAACACGACCATGCGGCGTGAAGTTTAATGGAAGGTCG 1821
Db 361 ATGGCCAGACAGGAAGTGGCAAGACACATATATGGCGGAGACCTCTCTGGGAAGGCC 420
QY 1822 AGGACTGCAGAACGGGCTCTACGCCATGGCGGCAAGGATGCTTTGTGACCTGATTA 1881
Db 421 AGAATGCATCCAAAGGATCTTATGCCATGGGCTCCCGGAGCGTCTTCTCTGAAGATC 480
QY 1882 TGCCCGGTTACCGCGCCATGAATCTAGTCTGCTCGGCGAGTTCTTTTGAGATTTACAGTG 1941
Db 481 AACCTGCTACCGGAAGTTGGGCTTGAAGTCTATGTGACATTTCTTCGAGATCTACAATG 540
QY 1942 GCAAGGCTTCGATCTTCTGTCCGACAGCAGAACTGCGCGTCTCTGGAGATGGTAAAC 2001
Db 541 GGAAGCTGTTTGACCTCTCAACAAAGAGGCCAAGCTGCGCGTCTGGAGGATGGCAAGC 600
QY 2002 AGCAAGTGCAGGTGGTGGGACTCACCGAGAGGTTGTCGATGGCGTCTGAGGAGTACTGA 2061
Db 601 AACAGGTGCAAGTGGTGGGCTGCGAGGAGATCTGGTTAACTCTGCTGATGATGTCATCA 660
QY 2062 AGCTCATCAGCAGCGCAATGCTGCCGCAACATCCGCGCAGACGCTCGGCCAACTCCAAT 2121
Db 661 AGATGATCGACATGGGCGAGCGCTGCAAGACCTCTGGCGAGACATTTGCCAACTCCAAT 720
QY 2122 CGTCGCTTCGACCGCTTTTTCAGATTTGCTCGCGCGGCGCAGGCTCGACGAGATCC 2181
Db 721 CCTCCGCTCCCAAGCTGCTTCCAAATTAATCTTTCGAGCTAAAGG-----GAGAATGC 774
```

QY 2182 ATGCAAGTTCCTGTTTCATGATCTGGGGGCAATGAGCGGGGGTGGGACATCTTCCTCGG 2241  
Db 775 ATGGCAAGTTCCTCTTTGGTAGATCTGGCAGGGAATGAGCGGGCGGACATCTCCAGTG 834  
QY 2242 CCATCGCGCAGACGCGTATGAGGGTCCGAGATTAAACAATCGTCTCGCCCTCAAGG 2301  
Db 835 CTGACCGCAGACCGCATGGAGGGCGCAGAAATCAACAAGATCTCTTAGCCCTGAGG 894  
QY 2302 AGTGCAATTCGTGCTGGGCAACAGTTCGCGCCCACTTTCGCCCTTCCTCCAAACTCA 2361  
Db 895 AGTGCAATTCGTGCTGGGCAACAGTTCGCGCCCACTTTCGCCCTTCCTCCAAACTCA 2361  
QY 2362 CCCAGGTGCTCGCGACTCTCTTCAATTCGCGAGAGCAGCAGATGATGATGAGCAGTGA 2421  
Db 955 CACAGGTGCTGAGGGACTCTCTTCAATTCGCGAGAGCAGCAGATGATGATGAGCAGTGA 2421  
QY 2422 TCTCGCGCGGACTTAGTCTCTCGCGAGCAGCAGCTCAACAGCTCGCGCTATGCGGATCGTG 2481  
Db 1015 TCTCACCAGGCATAAGCTCTCTGTAATATATCTTTAAACACCTTGAGATATGACAGGG 2481  
QY 2482 TCAAGGAGCTG 2492  
Db 1075 TCAAGGAGCTG 1085

RESULT 11  
US-09-594-669-15  
; Sequence 15, Application US/09594669  
; Patent No. 6331424  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Sakowicz, Roman  
; TITLE OF INVENTION: No. 6331424el motor proteins and methods for  
; FILE REFERENCE: 1042  
; CURRENT APPLICATION NUMBER: US/09/594,669  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: US 09/295,612  
; PRIOR FILING DATE: 1999-04-20  
; PRIOR APPLICATION NUMBER: US 09/314,464  
; PRIOR FILING DATE: 1999-05-18  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15  
; LENGTH: 2740  
; TYPE: DNA  
; ORGANISM: Human  
US-09-594-669-15

Query Match 14.6%; Score 449.8; DB 3; Length 2740;  
Best Local Similarity 62.4%; Pred. No. 1.5e-105;  
Matches 744; Conservative 0; Mismatches 437; Indels 12; Gaps 2;  
QY 1300 TGAAGAGCTGGAGCAGCTGAGGAGAAATCGCGAGAGCGACGCGCCGACAGCGGAG 1359  
Db 638 TGAAGAGCTGGAGGAAATGAGGAGAAATGAGGAGAGAGAGAGAGAGAGAGAGAGAG 697  
QY 1360 TGAAGAGAGAGAGAGTGGCGCTGATGAACAGGATCCGGGCAATCCAACTGGGAGCG 1419  
Db 698 TGAGATGAAGAGAG-----CTCAGGAGTATGACAGTAGTCTTCCAACTGGGATTTG 751  
QY 1420 CGCAATATGATCGCGAATATCAGAGCAGCTGGAATTTGTGCGGTGCTCGATGGCCAGG 1479  
Db 752 CCGCAATGATTAAGAAATTCGGGCTACTTTGGAATGTCATCCACTTACTATGATGATC 811  
QY 1480 CGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1539  
Db 812 CTATCGAAGAGCAGAGATATGATGATGATGATGATGATGATGATGATGATGATGATG 871  
QY 1540 TCAATCGAAGAGAGTTCGATGCTATTTTCGCTGCGCGAGAGCAGATGCTCATCTGTCAG 1599  
Db 872 TGGCCAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 931

QY 1600 AGCGCGCAGCAAGGTCGACCTCACCAGTTCCTGGAGAACCAAGTTCCTGCTTCGACT 1659  
Db 932 AACCCAAAGTTGAAGTGGACTTAACAAGATATCTGGAGAACCAAGCATTCCTGCTTCGACT 991  
QY 1660 ACSCCTTCAACGACACGTCGCAACATGCCATGATATACAAATACAGCAGCAAGCGCTGG 1719  
Db 992 TTGCATTTGATGAAACAGCTTCGAATGAAGTTGTCTACAGGTTTACAGCAGGCGCACTGG 1051  
QY 1720 TGAACCACTTTTCGAGGGCGGAATGGCGACGTCCTTCGCTACGGCCAGCAGCGGATCGG 1779  
Db 1052 TACAGACATCTTTGAGAGTGGAAACAACTTGTGTTGATATGCGCCAGCAGGAAAGTG 1111  
QY 1780 GCAAAACGACACATCGGCGGTGAGTTTAAATGGAAGGTGCAAGGCTGCAAGAACCGCA 1839  
Db 1112 GCAAGACACATATATGGCGGAGACCTCTCTCTGGAAAGCCAGAAATGTCATCCAAAGGGA 1171  
QY 1840 TCTACCGCATGGCGGCAAGGATGCTTTGTGACCTCTGATATATGCGCGTTACCGCGCCA 1899  
Db 1172 TCTATGCCATGGCGCTCCCGGAGCTCTTCTCTCTGGAAGATCAACCTCTGCTACCGGAAGT 1231  
QY 1900 TGAATCTAGTCTGCTCGGCCAGTTCCTTTGAGATTTTACAGTGGCAAGGCTCTTCGATCTTC 1959  
Db 1232 TGGGCTTGAAGTCTATGACATCTTTCGAGATCTACATGGGAAGCTGTTTGACCTGC 1291  
QY 1960 TGTCCGACAAAGCAGAACTGCGCGTCTCGAGGATGTTAAACAGCAAGTSCAGGTGGTG 2019  
Db 1292 TCAACAAAGAAAGCCAAAGCTGCGCGTCTGGAGGAGCGCAAGCAAGGTGCAAGTGGTG 1351  
QY 2020 GACTCACCAGAAAGTGGTTCGATGGCGTTCGAGGAGTACTGAAGCTCATCCAGCAGCGCA 2079  
Db 1352 GGTTCGAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGCTCGACATGGGCA 1411  
QY 2080 ATGTCGCCCGAAACATCCGCGCAGAGCTCGCGCAACTTCCAAATTCGTCGCTTCGCGCG 2139  
Db 1412 GCGCTGCGAGAACTCTGGGAGAGCATTTGCCAACTCCAAATTCCTCCCGCTCCACGCGT 1471  
QY 2140 TTTTCCAGATTTGCTGCGCGCGCAGGCTCGACGAGATCCATGGCAAGTTCCTCGTTCA 2199  
Db 1472 GCTTCCAAATTTTCTTCGAGCTAAAG-----GAGATGCAATGCAAGTTCCTCTTTGG 1525  
QY 2200 TCGATCTGGCGGCAATGAGCGGGGGTGGACACTTCTCGCGCGATTCGGCAGACGCGTA 2259  
Db 1526 TAGATCTGGCAGGAAATGAGCGGCGCAGACACTTCCAGTGTGTGACCGGAGCGCGCA 1585  
QY 2260 TGGAGGTGCGCAGATTAACAATCGCTCTCGCCCTCAAGGAGTGCATTCGTCGCTGG 2319  
Db 1586 TGGAGGCGCAGAAATCAACAAGATCTCTTAGCCCTGAAGGAGTGCATCAGGGCCCTGG 1645  
QY 2320 GCAACAGTCCGCCCACTTTGCCCTTCCTGCTCTCCAACTCACCAGGTGCTGCGGACT 2379  
Db 1646 GACAGAACAAAGCTCAGACCCCGTTCCGTCGAGAGCAAGCTGACACAGGTGCTGAGGACT 1705  
QY 2380 CTTTCATTTGGCAGAGCAAGAGCTGTCATGATAGCCATGATCTCGCGGAGCTTAGCT 2439  
Db 1706 CTTTCATTTGGGAGAACTCTAGGACTTGCATGATTTGCCACGATCTCACCAGGCATAGCT 1765  
QY 2440 CTTGCGAGCAGACGCTCAACAGCTGCGCTATGCGGATGCTGTCAGGAGCTG 2492  
Db 1766 CTTGGAATATATCTTTAAACACCTTGAGATATGACAGAGGTCAAGGAGCTG 1818

RESULT 12  
US-09-595-684B-32  
; Sequence 32, Application US/09595684B  
; Patent No. 6544766  
; GENERAL INFORMATION:  
; APPLICANT: Beraud, Christophe  
; APPLICANT: Ohashi, Cara  
; APPLICANT: Sakowicz, Roman  
; APPLICANT: Vaisberg, Eugeni  
; APPLICANT: Wood, Kenneth  
; APPLICANT: Yu, Ming  
; TITLE OF INVENTION: Human kinesins and methods of producing

```
/ TITLE OF INVENTION: and purifying human kinesins
/ FILE REFERENCE: cytopo36
/ CURRENT APPLICATION NUMBER: US/09/595,684B
/ CURRENT FILING DATE: 2002-06-24
/ PRIOR APPLICATION NUMBER: 09/295,612
/ PRIOR FILING DATE: 2000-04-20
/ NUMBER OF SEQ ID NOS: 105
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 32
/ LENGTH: 2740
/ TYPE: DNA
/ ORGANISM: Human
/ US-09-595-684B-32

Query Match          14.6%; Score 449.8; DB 3; Length 2740;
Best Local Similarity 62.4%; Pred. No. 1.5e-105;
Matches 744; Conservative 0; Mismatches 437; Indels 12; Gaps 2;

QY 1300 TGAAGAGGTGGAGCGACTGAAGGAGAATCGCGAGAAGCGAGCGCGCGGACAGAGCGCGGAGG 1359
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
638 TGAAGGAACTGGAAGAAATGAAGAACAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 697
QY 1360 TGAAGGAGAGAGGTGGCGCTGATGAACAGAGATCCGGGCAATCCAACTGGGAGAGCGG 1419
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
698 TGAGAAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTATTTCCAAACTGGGAATTG 751
QY 1420 CGCAAATGATACGGAATATCAGAGCAGCGTGGAAATTTGTGCGGCTGCTCGATGCCAGG 1479
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
752 CCGGATGATTAAGAAATTCGGGCTACTTTGGAATGTCATCCACTTACTATGATGATC 811
QY 1480 CGTCGATGACCATCAGATCAGATGTGCGTGGCAAGCGTCCCATTAGCGCGCAAGGAGG 1539
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
812 CTATCGAAGAGCACAGAAATATGTGCTGTGTTAGGAACGCCACTGATTAAGCAAGAT 871
QY 1540 TCAATCGCAGAGAGATCGATGTCATTCGTGCGCGCGCAAGAGCATGCTCATGCGACG 1599
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
872 TGGCCAAAGAAAGAAATGATGTGATTTCCATTCTCTAGCAAGTGTCTCTCTTGGTACATG 931
QY 1600 AGCGCGCGCAGCAAGGTCCGACCTCACCAAGTTCTCGGAGAACACAAAGTTTCGCTTCGACT 1659
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
932 AACCCAAAGTTGAAAGTGAACTTAACAAAGATATCTGGAGAACCAAGCATTCGCTTTGACT 991
QY 1660 AGCGCTTCAACGACACGTGCGCAATGCCATGGTATACAAATACACAGCCCAAGCGCTGG 1719
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
992 TTGCATTTGATGAACACAGCTTCGAATGAAGTTGCTACAGGTTACACAGAGGCCACTGG 1051
QY 1720 TGAACACCATTTTCAGGCGGGAATGGCGAGTGCTTCGCTACGCGCAGAGGGGATCGG 1779
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1052 TACAGACAATCTTTGAAGGTGGAAAGCAACTTGTGTTGTCATATGCGCCAGACAGGAAGTG 1111
QY 1780 GCMAACCGCACACCATGGCGGTGAGTTTAATGGAAGGTGCAGGACTGCAGAACCGCA 1839
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1112 GCAAGACACATATATGGCGGAGACCTCTCTGGGAAGCCGAGATGCAATCCAAAGGGA 1171
QY 1840 TCTACGCCATGGCGGCCAAGGATGCTTTGTGACCCCTGAATATGCGCGTTTACCGCGCCA 1899
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1172 TCTATGCCATGGCTCTCCGGGAGCTTCTCTCTGGAAGATCAACCTCTGCTACCGGAGT 1231
QY 1900 TGAATCTAGTCTGCTCGGCCAGTTCTTTTGGATTTACAGTGGAAGGTCTTCGATCTTC 1959
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1232 TGGGCTTGGAAAGTCTATGTGACATTTCTCGAGATCTACAATGGGAAGCTGTTTGACCTGC 1291
QY 1960 TGTCCGACAGCAGAAACTGCGCTCTCGGAGGATGTTAAACAGCAAGTGCAGTGGTGG 2019
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1292 TCAACAGAGAGCCAGCTGCGGTCTGGAGGAGCGGCAAGCAAGTGCAGTGGTGGTGG 1351
QY 2020 GACTCACCGAGAAGGTGGTTCGATGGCGTCGAGGAGGTACTGAAGCTCATCCAGCACGGCA 2079
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1352 GGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGCTGCACATGGGCA 1411
QY 2080 ATGTGCGCCGAAACATCCGCGCAGAGGTGGCGCAACTCCAAATTCGTGGGTTTCGACGCGG 2139
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1412 CGGCTTGAGAACTCTGGGAGAGACATTTGCGCAACTCCAAATTCCTCCGCTCCACCGCT 1471
```

```
QY 2140 TTTTCCAGATTGTGTGCGCGCGCAGGCTCGACGAGATCCATGGCAAGTTCCTGCTTCA 2199
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1472 GCTTCCAAATTAATCTTCGAGCTAAAG-----GAGAAATCGATGGCAAGTTCCTTTGG 1525
QY 2200 TCGATCTGCGCGGCAATGAGCGGCGGTGGACACTTCTCGGCGGATCGGCGAGACGCGTA 2259
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1526 TAGATCTGGCAGGGAATGAGCGGCGGAGACACTTCCAGTGTGACCGGCGAGACCGCA 1585
QY 2260 TGGAGGGTGGCGAGATTAACAATCGTGTGCGGCTTCAAGAGTGCATTCGTGCTGG 2319
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1586 TGGAGGGCGCAGAAATCAACAAGAGTCTCTTAGCCCTGAAGAGTGCATCAGGCGCCTGG 1645
QY 2320 GCAAAACAGTCCGCCCACTTTCCTTCCGCTCTCCAACTCACCAGGTGCTGCGGACT 2379
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1646 GACAGAAACAGGCTCACACCCCTTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGACT 1705
QY 2380 CGTTCAATGGCGAAGAGCAAGACGTCGATGATAGCATATCTCGCGGAGACTTAGCT 2439
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1706 CTTTCAATGGGAGAACTCTAGGACTTGCATGATGCGCAGATCTCACCAGGCATAGCT 1765
QY 2440 CCTCGAGCACACGCTCAACACGCTGCGCTATGCGGATCGGTCAAGGAGCTG 2492
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1766 CCTGTGAATATATCTTTAAACACCTGAGATATGACAGAGGTCAAGGAGCTG 1818

RESULT 13
US-10-112-432-15
; Sequence 15, Application US/10112432
; Patent No. 6638754
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: No. 6638754el motor proteins and methods for
; FILE REFERENCE: 1042A
; CURRENT APPLICATION NUMBER: US/10/112,432
; CURRENT FILING DATE: 2002-03-29
; PRIOR APPLICATION NUMBER: 09/724,215
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: 09/314,464
; PRIOR FILING DATE: 1999-05-18
; PRIOR APPLICATION NUMBER: 09/594,669
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 2740
; TYPE: DNA
; ORGANISM: Human
US-10-112-432-15
```

```
Query Match          14.6%; Score 449.8; DB 3; Length 2740;
Best Local Similarity 62.4%; Pred. No. 1.5e-105;
Matches 744; Conservative 0; Mismatches 437; Indels 12; Gaps 2;

QY 1300 TGAAGAGGTGGAGCGACTGAAGGAGAATCGCGAGAAGCGAGCGCGCGGACAGCGCGGAG 1359
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
638 TGAAGGAACTGGAAGAAATGAAGAACAGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 697
QY 1360 TGAAGGAGGAGAGGTGGCGCTGATGAACAGGATCGGGCAATCCAACTGGGAGAGCGG 1419
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
698 TGAGAAATGAAGAGAG-----CTCAGGAGTATGACAGTAGTATTTTCCAAACTGGGAATTG 751
QY 1420 CGCAAATGATACGGAATATCAGAGCAGCGTGGAAATTTGTGCGGCTGCTCGATGGCCAGG 1479
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
752 CCGAATGATTAAGAAATTTCCGGCTACTTTTGGAAATGTCATCCACTTACTGACTGATC 811
QY 1480 CCGTCGATGACCATCAGATCAGATGTGCTGCGCAAGCGTCCCAATTCGCGGAGGAGG 1539
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
812 CTATCGAAGAGCACAGAAATATGCTGTGTTAGGAAACGCCCACTGAATTAAGCAAGAT 871
```

Qy	1540	TCAATCGCAAGAGATCGATGTCATTTTCGGTCCGCGCAAGGACATGCTCAATCGTGCAAG	1599
Db	872	TGGCCAAAGAAAGAAATTAATGATGTAATTTCCATTTCTAGCAAGTGTCTCTCTTTGGTACATG	931
Qy	1600	AGCCGCGCAGCAAGGTGACCTCACCAAGTTCTCTGSGAAGCCACAAGTTTTCGCTTCGACT	1659
Db	932	AACCCAGTTGAAAGTGGACTTAACAAAGTATCTGGAGAACCAAGCAATTTCTCTTTGACT	991
Qy	1660	ACGCTTCTAACGACACGCTGGCGCAATGCGTATGATGTTATCAAAATACACAGCCAAAGCGTGTGG	1719
Db	992	TTGCATTTGATGAACAACAGCTTCGGAATGAAGTTGTCTACAGTTTCACAGCAAGGCCACTGG	1051
Qy	1720	TGAAAAACAATTTTCGAGGGCGGAATGGGACGTCCTTCGCTACGGCCAGACGGGATCGG	1779
Db	1052	TACAGACAAATCTTTGAAGGTTGAAAAGCAACTCTGTTTGTATATGCGGCACAGCAAGAAATG	1111
Qy	1780	GCAAAAACGACACACCATGGCGGTGAGTTTAATGAAAAGGTGCAGACTGTCAAAGAAACGCA	1839
Db	1112	GCAAGACACATATATATGGCGGAGACCTCTCTGGGAAAGCCAGAAATGCAATGCCAAGGGA	1171
Qy	1840	TCTACGCCATGGCGGCCAAAGGATCTTTTGTGACCTCTGAATATATGCCGCTTACCGCGCCA	1899
Db	1172	TCTATGCCATGGCTCCCGGAGCGTCTTCTCTCTGAAGAACTCAACCTGCTACCGGAAGT	1231
Qy	1900	TGAATCTAGTCGTCTCGCCAGTTTCTTTGAGATTTACAGTGGCAAGGTCTTCGATCTTC	1959
Db	1232	TGGCCCTGGAAATCTATGTGACATCTCTTCAGAGATCTACAAATGGGAAAGCTGTTTGACCTGC	1291
Qy	1960	TGTCGCGCAAGCAGAACTCGCGCGTCTGGAGGATGTTAAACAGCAAGTGCAGTGTGTGG	2019
Db	1292	TCAACAGAAGGCCAAGCTGCGCGTGTGTGAGGACGGCAAGCAACAGTGCAGTGTGTGG	1351
Qy	2020	GACTCACCGAAGGTGGTCGATGGCGCTCGAGGAGGTACTGAAGCTCATCCAGCACGCGCA	2079
Db	1352	GGCTGCAGGAGCATCTGGTTAACTCTGCTGATGATGTCATCAAGATGCTCGACATGGCA	1411
Qy	2080	ATGCTGCGCGNAACATCGGCGCAGACGTGCGGCAACTCCAATTCGTGCGGTTTCGACGCGG	2139
Db	1412	GCGCCTCGAACAACCTCTGGGCAGACATTTGCAACTCTCAATTCCTCCGCTCCCAACGCGT	1471
Qy	2140	TTTTTCCAGATTGCTCGCGCGCAGAGGCTCGACGAAGATCCATGGCAAGTCTTCGTTCA	2199
Db	1472	GCITTCAAAATTATCTTCGAGCTTAAAG-----GAGAAATGCATGGCAAGTCTCTTTGG	1525
Qy	2200	TCGATCTGGCGGGCAATGAGCGGGCGGTGGACACTTCTCGCGCGGATCGGCAGACGCGTA	2259
Db	1526	TAGATCTGGCAGGGAATGAGCGAGCGCAGACACTTCCAGTGTGTGACCGGCAGACCGCA	1585
Qy	2260	TGGAGGTGCGGAGATTAAACAAATTCGTGCTGGCCCTCAAGAGTGCATTCGTGCGGTGG	2319
Db	1586	TGGAGGCGCGAGAAATCAACAAGAGTCTCTTAGCCCTTGAAGGAGTGCATCAGGGCCCTCG	1645
Qy	2320	GCAACAGTGGGCCCACTTGCCCTTCGCTGTCTCCAAACTCACCCAGGTGCTCGGGACT	2379
Db	1646	GACAGAACAGGCTCAACACCCGTTTCGTGAGAGCAAGCTGACACAGGTGCTGAGGACT	1705
Qy	2380	CGTTTATTGGCGAAGACGAACGCTGCATGATAGCCATGATCTCGCCGGGACTTAGCT	2439
Db	1706	CCTTCAATTGGGAGAACTCTAGGACTTTGCATGATTGCCACGATCTCACAGGCATAAGCT	1765
Qy	2440	CTTGCAGACACGCTCAACACGCTGCGCTATATCGGATTCGTGTCAAGGAGCTG	2492
Db	1766	CTGTGTAATATACCTTTAAACACCTCTGAGATATGCAGACAGGGTCAAGGAGCTG	1818

## RESULT 14

RESOLUTION 14  
US-09-849-602-5  
; Sequence 5, Application US/09849602  
; Patent No. 6794501  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew J.  
; APPLICANT: Old, Lloyd J.

Qy	2140	TTTTCCAGATTGTGTCGGCCGCAGGGCTCGACAAAGTCATGGCAAGTCTCTGTTCA	2199
Db	1472	GCTTCCAAATTAATCTTCGAGCTAAAG-----GAGAAATGCATGGCAAGTCTCTTTGG	1525
Qy	2200	TCGATCTCTGGCGGGCAATAGACGGGGCTGGACACTTCTCTGGCCGATCGGCAGACGGTA	2259
Db	1526	TAGATCTTGGCAGGGAATAGCGAGCGCAGACACTTCCTAGTGTCTGACCGGCAGACCCGCA	1585
Qy	2260	TGGAGGTCGCGAGATTAAACAAATCGCTGTGGCCCTCAAGGAGTGCAITCTGTCGCTGG	2319
Db	1586	TGGAGGGCGCAAAATCAACAAGAGTCTCTTAGCCCTGAAGGAGTGATCAGGGCCCTGG	1645
Qy	2320	GCAACAGTCGGCCCACTTGCCCTTCGCTGTCTCAAACTCACCCAGTGCTCGCGACT	2379
Db	1646	GACAGACAGGCTCACACCCGCTTCGTCGAGACCAAGCTGCACAGTCTCTAGGGACT	1705
Qy	2380	CGTTCAATTGGCGAAGACGATGTCATGATAGCCATGATCTCCCGGGACTTAGCT	2439
Db	1706	CTTTCATTGGGAGAACTTAGGACTTGCATGATTGCCAGATCTACAGGCATTAAGCT	1765
Qy	2440	CCTCGAGCACAGCTCAACACGTCGCTATGCGGATCTGTCTCAAGGAGCTG	2492
Db	1766	CTGTGGAATACTTTAAACACCTCGAGATATGCAGACAGGGTCAAGGAGCTG	1818

```

RESULT 15
US-09-594-669-5
; Sequence 5, Application US/09594669
; Patent No. 6331424
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Sakowicz, Roman
; TITLE OF INVENTION: NO. 6331424el motor proteins and methods for
; TITLE OF INVENTION: their use
; FILE REFERENCE: 1042
; CURRENT APPLICATION NUMBER: US/09/594,669
; CURRENT FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/295,612
; PRIOR FILING DATE: 1999-04-20
; PRIOR APPLICATION NUMBER: US 09/314,464
; PRIOR FILING DATE: 1999-05-18
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1088
; TYPE: DNA
; ORGANISM: Human
US-09-594-669-5

```

Query Match	13.8%;	Score 425.6;	DB 3;	Length 1088;
Best Local Similarity	65.2%;	Pred. No. 1.7e-99;		
Matches 645;	Conservative 0;	Mismatches 339;	Indels 6;	Gaps 1;
QY	1503	GTGTGGTGGCAAGCGTCCCATTTAGCGGCACAGAGGTCAATCCCAAGGAGATCGATGTC	1562	
DB	15	GTCTGTGTTAGGAACGCCCACTGAATAAGCAAGATTTGGCCAAAGAAGAAATTTGATGTG	74	
QY	1563	ATTTTCGGTGGCGGCAAGGACATGCTCATCGTGCACAGCGCGGCACCAAGGTGCACTC	1622	
DB	75	ATTTTCNATCTTAGCAAGTGTCTCTCTTGGTATACGAACCCAAAGTTGAAGTGGACTTA	134	
QY	1623	ACCAAGTTCCTGGAGAACCAAGATTTTCGGTTTCGACTACGCTTCAACGACACGCTGGCAC	1682	
DB	135	ACAAAGTATCTGGAGAACCAAGCATTTCTGTTTGAACCTTTGCATTTGATGAACACGCTTCG	194	
QY	1683	AATGCCATGTTATACAAATACACAGCCAAGCCGTTGGTGTGAAAAACATTTTCGAGGGCGGA	1742	
DB	195	AATGAAGTTGCTACAGGTTTCACAGCAAGGCCACTGGTACAGACAATCTTTGAAGGTGGA	254	
QY	1743	ATGGCAGCTGCTTTCGCTTAGCGCTACGGGCAGACGGGATTCGGGGCAAAACGACACATCGGGCGGT	1802	
DB	255	AAAGCAACTGTGTTTTCGATATGGCCACACAGGAAGTGGCAAGACACATATATGGGCGGA	314	

QY	1803	GAGTTTAAATGAAAGGTCGAGGACTGCAAGAACGGCATCTACGCCATGGCGGCCAAGGAT	1862
Db	315	GACCTCTCTGGAAAGCCAGAAATGCATCAAAAGGGATCTATGCCATGGCTCCCGGGAC	374
QY	1863	GTCTTTGTGACCCCTGAATATGCGCGGTATCGCGGCCATGAATCTAGTCTGCTCGGCCAGT	1922
Db	375	GTCTTCCTCTCTGAAGAAATCAACCCCTGCTACCGGAAGTTGGCGCTTGAAGTCTATGTGACA	434
QY	1923	TTCTTTTGAGATTATACGTGGCGAAGTCTTCGATCTTCTGTCGCAACAGCAGAAACTGCGC	1982
Db	435	TTCTTCGAGATCTACAAATGGGAAGCTGTTTGAACCTGCTCAACAAGAAAGCCAAAGCTGCGC	494
QY	1983	GTCTGGAGGATGGTAAACACGAAAGTCGAGTGGTGGGACTCACCGAGAAGGTGGTCTGAT	2042
Db	495	GTGCTGGAGGATGGCAAGCAACAGGTGCAAGTGTGGGGCTGCAGGAGCATCTGTTTAAAC	554
QY	2043	GGGCTCGAGAGGATCTGAAGCTCATCCAGCAGCGCAATGCTGCCGGAACATCCGGCCAG	2102
Db	555	TCTCTGATGATGTTCATCAAGATGATCGAATGGCGAGCGCTGCAGAACCTCTGGGCAG	614
QY	2103	ACGTGCGCCAACTCCAATTCGTGCGGTTCCAGCGCCGTTTTCCAGATTGCTGTCGGCCG	2162
Db	615	ACATTTGCCAACTCCAATTCCTCCCGCTCCACGCGTGCTTCCAAATTAATCTTCAGACT	674
QY	2163	CAGGGCTCGACGAGATGCCATGGCAAGTTCTTCGTTTCATTCGATCTGGCGGGCAATGAGCGG	2222
Db	675	AAAGG-----GAGATGCAATGGCAAGTTCTCTTTGGTAGATCTGGCAGGGAAATGAGCGA	728
QY	2223	GGCGTGACACTTCTCGCGCCGATCGGCAGACGCGTATGGAGGGTCCGAGATTAAACAA	2282
Db	729	GGCGCGGACACTTCCAGTGTGTACCGCGCAGACCCGATGGAGGGCGCAGAAATCAACAAG	788
QY	2283	TCGCTGTGCGCCCTCAAGAGTGCATCTGTGCGTTGGGCAACAGCTGCGCCCACTTGCCC	2342
Db	789	AGTCTCTTAGCCCTGAAGGAGTGCATCAGGGCCCTCGGACAGAAACAGGGCTCACACCCCG	848
QY	2343	TTCCGTGTCTCCAAACTCACCCAGGTGCTGCGGACTTCGTTTCATTCGCGCAGAGAAGCAAG	2402
Db	849	TTCCGTGAGAGCAAGCTGACACAGGTGCTGAGGGAATCTCTTCAATGGGGAGAAATCTTAGG	908
QY	2403	ACGTGCATGATGCCATGATCTCCCGGGACTTTAGTCTCTCTGCGAGCACACGCTCAACACG	2462
Db	909	ACTTGCATGATTGCCACGATCTCACCAGGCATAAGCTCCTGTGAATATATCTTTAAACACC	968
QY	2463	CTGGCTATGCGGAATCGTGTCAAGAGACTGG	2492
Db	969	CTGAGATATGCACACAGGGTCAAGGAGCTG	998

Search completed: May 12, 2006, 21:18:16  
Job time : 532 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 12, 2006, 17:35:46 ; Search time 11790 Seconds  
(without alignments)  
12226.548 Million cell updates/sec

Title: US-10-840-060-142

Perfect score: 3081

Sequence: 1 aaactaaaaatgtgttc.....gccgaatggcaacgcttagt 3081

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

1: gb\_est1.\*  
2: gb\_est2.\*  
3: gb\_est3.\*  
4: gb\_hc.\*  
5: gb\_est4.\*  
6: gb\_est5.\*  
7: gb\_est6.\*  
8: gb\_est7.\*  
9: gb\_gss1.\*  
10: gb\_gss2.\*  
11: gb\_gss3.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	917	29.8	1101	10	CNS0011P
2	761.6	24.7	779	2	BF499568
3	752.6	24.4	759	2	BF484448
4	750.2	24.3	763	2	BF498461
5	720.4	23.4	722	2	BF489668
6	718	23.3	1105	8	DR122830
7	692	22.5	1153	8	DR135726
8	685.8	22.3	720	2	BF498964
9	683.6	22.2	702	2	EG633810
10	680.6	22.1	687	2	BF495560
11	678	22.0	686	2	BF499204
12	674.4	21.9	689	3	BI635599
13	663.8	21.5	679	2	BF495237
14	663.6	21.5	670	2	EG634444
15	662.4	21.5	664	3	BI603735
16	659.8	21.4	744	2	BF505411
17	659	21.4	660	3	BI592847
18	656	21.3	656	1	AI259349
19	641	20.8	641	2	BF488088
20	629	20.4	629	1	AI388100
21	625.2	20.3	630	2	BF503725
22	619.6	20.1	626	2	BF491708

23	605.8	19.7	719	1	AI546445	AI546445	LD48021.5
24	599	19.4	651	3	BI372858	RS60047.5	
25	591.6	19.2	598	2	BF506249	AT08677.5	
26	590.4	19.2	695	1	AA949737	LD29734.5	
27	589.2	19.1	605	3	BI354473	GM26945.5	
28	589	19.1	740	1	AA952135	LD29208.5	
29	588.4	19.1	682	2	BF488658	AT24155.5	
30	587.8	19.1	591	2	BF492432	LD29373.5	
31	582.6	18.9	749	2	BF506266	AT08703.5	
32	582.2	18.9	642	1	AI546390	LD47952.5	
33	577	18.7	586	1	AI532286	SD03743.5	
34	575.2	18.7	721	7	CK660542	LP19179.5	
35	575	18.7	575	1	AI515654	AI515654	LD47558.5
36	574	18.6	574	3	BI609578	RH14378.5	
37	574	18.6	574	3	BI609874	RH14717.5	
38	574	18.6	575	3	BI609406	RH14181.5	
39	573.2	18.6	587	1	AI533352	SD05134.5	
40	571.8	18.6	587	2	BF492824	AT01227.5	
C	41	570	664	10	AG924280	AG924280	Drosophila
	42	568	18.4	568	2	BF491744	AT28556.5
	43	563.4	18.3	565	3	BI609731	RH14552.5
	44	560.8	18.2	572	3	BI610275	BI15384.5
	45	558.4	18.1	597	2	BE975372	BE975372

#### ALIGNMENTS

RESULT 1  
CNS0011P 1101 bp DNA linear GSS 03-JUN-1999  
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #  
DEFINITION BACR04G01 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.  
ACCESSION AL061056 GI:4939519  
VERSION AL061056.1  
KEYWORDS GSS.  
ORGANISM Drosophila melanogaster (fruit fly)  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
REFERENCE 1 (bases 1 to 1101)  
AUTHORS Direct Submission  
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :  
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
COMMENT - Web : www.genoscope.cns.fr)  
Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see [http://www.fruitfly.org/TheBDGP/Drosophila melanogaster BAC library](http://www.fruitfly.org/TheBDGP/Drosophila_melanogaster_BAC_library) was prepared by Kazutoyo Ooeagawa and Aaron Mammose in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp. The same strain used for the BDGP's PI and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

Location/Qualifiers  
1. .1101  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACR04G01"  
/clone\_lib="RPCI-98"  
/notes="end : TET3"

FEATURES  
source

ORIGIN



Query Match		29.8%; Score 917; DB 10; Length 1101;
Best Local Similarity		90.9%; Pred. No. 4e-227;
Matches		957; Conservative 41; Mismatches 52; Indels 3; Gaps 3;
QY	269	CACACACACACACTGGGACATCGAAAAAACAACAGACGCTGTGCGGACCTCGA 328
DB	5	CACACACACACTTAGAACACATCAAAAAAACAACAAACGCTGTCGGGACCTCAA 64
QY	329	CAAGCAGTGGGTTCGATTAGTTGTCAATGCGCTTGAATTCGGTTGGGCTTAGTTCC 388
DB	65	CNAACATTTTKYTCATTTATTTTTCATGCGCTTAAATTCGTTCGGGCTTATTTTC 124
QY	389	ACAAGTTTATCGCTCGTCAAGAAACAACAGAAATAAATTTTCGACCTAAAAATCTG 448
DB	125	ACAATTTTATCGCTCTTCAAAAAACAACAAATAAATTTATTTTCGACCTAAAAATCTA 184
QY	449	ACTAAATGTGTTTTTGTATTATGTTATTTATTTAGGCACATTTTGCACACCAACGTAG 508
DB	185	ACTAAATTTTTTTTTTTTTTTATKTATTTATTTAGGCACATTTTGCACACCAACGTA- 243
QY	509	TTACTACATCTACGACTAACGGAACCTCTCTCGAAGCAGTGAAGTTGCTGTCCATCAA 568
DB	244	TTACTACATCTACGACTTAAGGAACCTCTCTCTCGAAGCAGTGAAGTTGCTGTCCATCAA 303
QY	569	GCAGTACTCGAGTTAACGCAGGATAGCCGGGAGAAAGAGAGATCGGTGGAGAAT 628
DB	304	GCAGTACTCGAGTTAACGCAGGATAGCCGGGAGAAAGAGAGATCGGTGGAGAAT 363
QY	629	AGAGATATACAGTGGAGTCAAGAGAAAGGATCATGACATGATTACGTTGGGGCAGAG 688
DB	364	AGAGATATACAGTGGAGTCAAGAGAAAGGATCATGACATGATTACGTTGGGGCAGAG 423
QY	689	CGTCAAGATCAAGCGGACGATGCGCGCTCCATGCGCGTGGTGGGTGATCAACCA 748
DB	424	CGTCAAGATCAAGCGGACGATGCGCGCTCCATGCGCGTGGTGGGTGATCAACCA 483
QY	749	GTCGGGCAAGTGCATCACAGTTCGAATGTTACGAGCGCGGCAACGAGGCGCAAGGAGT 808
DB	484	GTCGGGCAAGTGCATCACAGTTCGAATGTTACGAGCGCGGCAACGAGGCGCAAGGAGT 543
QY	809	AGAACTGGACGCATATCAAGTTCGAGAGTAAATGCAAGATCTGTCGAAACGCA 868
DB	544	AGAACTGGACGCATATCAAGTTCGAGAGTAAATGCAAGATCTGTCGAAACGCA 603
QY	869	CGCGCCCGGAGCCCAAGAAACAAGCCACCGCGCGATCAACTCTCGCGTAAATCCAC 928
DB	604	CGCGCCCGGAGCCCAAGAAACAAGCCACCGCGCGATCAACTCTCGCGTAAATCCAC 663
QY	929	ACAATCGGCTATCGGTGGCAATCTCACAGCCGTATGACCATGCGCGGAAACATGCTGAA 988
DB	664	ACAATCGGCTATCGGTGGCAATCTCACAGCCGTATGACCATGCGCGGAAACATGCTGAA 723
QY	989	CAAGATCCAGGAAGCAGTCGATTCGCAATTCGATGTCAGCAGCAATAGCGTGAATC 1048
DB	724	CAAGATCCAGGAAGCAGTCGATTCGCAATTCGATGTCAGCAGCAATAGCGTGAATC 783
QY	1049	AAACAGCACTCCCACTACGCGCGCGAGGTGGTGGCAACCAACGTCGACGACCAC 1108
DB	784	AAAMAGCACTCCCACTACGCGCGCGAGGTGGTGGCA-CWCAACSTFCGACGACCAC 842
QY	1109	TGGATTACAGCGTCCAGGTACTCGAAGCTGTCTACCGGCGAGCAGACAGCAAGGATCGC 1168
DB	843	TGGATTACAGCGTCCAGGTACTCGAAGCTGTCTACCGGCGAGCAGACAGCAAGGATCGC 902
QY	1169	CTCGGCGGTGCTTAATAACAATTTGCCAATCCACGCGCGGACGACGAGTCTGGTCCGGC 1228
DB	903	CYCGGCGGTGCTTAATAACAATTTGCCAATCCACGCGCGGACGAGTCTGGTCCGGC 962
QY	1229	GGCACAAGGATCGCCATCGCGGCCAACCCAGCGGAGCTGGCGGCTAGTACCGGGC 1288
DB	963	SGCACAARGAGTGR-CMYKSGGCCAAWACCMMRGAGCTTTCGCGSCTATWACCGGCSA 1021
QY	1289	ATCGCACGCTTGAAGAGGTGGAGGCACTGAA 1321

DB	1022	TYCMCCATTWAAAAGGTGARCSTACTGAAGAA 1054
RESULT 2		
BF499568		779 bp mRNA linear EST 19-APR-2001
LOCUS		AT14153.Sprime AT Drosophila melanogaster adult testes pOTB7
DEFINITION		Drosophila melanogaster cDNA clone AT14153 5 similar to CGI453:
KEYWORDS		PBAN0001453 'motor protein' located on: X 10A6-10A7;; 04/09/2001, mRNA sequence.
ACCESSION		BF499568
VERSION		BF499568.2 GI:13691428
SOURCE		EST.
ORGANISM		Drosophila melanogaster (fruit fly)
REFERENCE		Drosophila melanogaster
AUTHORS		Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
TITLE		1 (bases 1 to 779)
JOURNAL		Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E., Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V., Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N., Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C. J., Nunoo, J., Pacieb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Celniker, S. and Rubin, G. M.
COMMENT		BDGP/HMI AT Drosophila EST Project Unpublished (2000) On Dec 6, 2000 this sequence version replaced gi:11582869. Contact: Stapleton, M. BDGP Lawrence Berkeley National Lab One Cyclotron Rd, Berkeley, CA 94720, USA Fax: 510 486 6798 Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu hit genomic AB003485: arm:X [10878159,11180508] estimated-cyto:10A3-10B11: 04/09/2001 Plate: AT.141 row: E column: 5 High quality sequence stop: 689. Location/Qualifiers 1. .779 /organism="Drosophila melanogaster" /mol_type="mRNA" /db_xref="taxon:7227" /clone="AT14153" /sex="male" /dev_stage="0-3 day old Ore-R males" /lab_host="Plates AT.10-AT.120: DHS-alpha. Plates AT.121-AT.319: DHS-alpha Tona" /clone_lib="AT Drosophila melanogaster adult testes pOTB7" /note="Organ: ADULT testes; Vector: pOTB7; Site: 1: Rcor1; Site: 2: Xho1; The mRNA for the testis library was made from testes and seminal vesicles hand dissected from 0-3 day old Ore-R males. RNA kindly provided by the lab of Margaret Fuller. Sized fractionated cDNAs were directly ligated into pOTB7. Plasmid cDNA library."
FEATURES		
source		
ORIGIN		
Query Match		24.7%; Score 761.6; DB 2; Length 779;
Best Local Similarity		99.4%; Pred. No. 1.1e-186;
Matches		775; Conservative 0; Mismatches 4; Indels 1; Gaps 1;
QY	387	CCACAAGTTTATCGTTCGTCAAGAAACAACGAATAAATAATTTTCGACCTAAAAATC 446
DB	1	CCACAAGTTTATCGTTCGTCAAGAAACAACGAATAAATAATTTTCGACCTAAAAATC 60
QY	447	TGACTAAATTTGTTTTTTTGTATTATTTAGGCACATTTTGCACACCAACCGT 506
DB	61	TGACTAAATTTGTTTTTTTGTATTATTTAGGCACATTTTGCACACCAACCGT 120
QY	507	AGTTACTACATCTACGACTAACCGAACTCCTCTCGAAGCAGTGGAGTGTGTCATC 566
DB	121	AGTTACTACATCTACGACTAACCGAACTCCTCTCGAAGCAGTGGAGTGTGTCATC 180



```
Db 661 CGCCGCGGAGTGTGGCACCACCAACGTCGACGACCACTGGATTACAGGTCCACGGT 720
QY 1129 ACTCGCAAGTGTACCGCGCAGCAGCAGACNAGGATCG 1167
Db 721 ACTCGCAAGTGTCTACCGCGCAGCAGCAGACNAGGATCG 759

RESULT 4
BF498461
LOCUS
DEFINITION
Ari2870.5prime AT Drosophila melanogaster adult testes pOTB7
Drosophila melanogaster cDNA clone AT12870 5 similar to CG12192:
FBan0012192 'motor protein' located on: 2R 59D1-59D1.; 04/09/2001,
mRNA sequence.
ACCESSION
BF498461
VERSION
BF498461.2 GI:13690358
KEYWORDS
EST.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 763)
REFERENCE
AUTHORS
Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Baxter,E.,
Berman,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V.,
Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N.,
Li,P., Liao,G., Miranda,A., Misra,S., Mungall,C.J., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C.,
Lewis,S.E., Celnikier,S. and Rubin,G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)
TITLE
BDGP/HMI AT Drosophila EST Project
JOURNAL
On Dec 6, 2000 this sequence version replaced gi:11581762.
COMMENT
Contact: Stapleton, M.
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
hit genomic AE003485: arm:X [10878159,11180508]
estimated-cyto:10A3-10B11: 04/09/2001
Plate: AT.128 row: F column: 10
High quality sequence stop: 673.
FEATURES
Location/Qualifiers
1..763
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT12870"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/notes="Organ: ADULT testes; Vector: pOTB7; Site 1: EcorI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vassicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."
ORIGIN
Query Match 24.3%; Score 750.2; DB 2; Length 763;
Best Local Similarity 99.0%; Pred. No. 1e-183;
Matches 755; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1090 CCACACGTCGACGACCACTGGATTACAGGTCACCGTACTCGCAAGCTGCTACCGGCC 1149
Db 1 CCACACGTCGACGACCACTGGATTACAGGTCACCGTACTCGCAAGCTGCTACCGGCC 60
QY 1150 AGCAGCAGACCAAGATCGCTCGCGGTGCTTAATACACATTGCCCAATCCACGCGCG 1209
Db 61 AGCAGCAGACCAAGATCGCTCGCGGTGCTTAATACACATTGCCCAATCCACGCGCG 120
```

```
QY 1210 CAGCCAGTGTCTGTCCGCGGCGCAACAGGAGTGCCTCAGCGCCACCAACCCAGGAGCTG 1269
Db 121 CAGCCAGTGTCTGTCCGCGGCGCACAGGGAGTGCCTCAGCGCCACCAACCCAGGAGCTG 180
QY 1270 GCGCGCTAGTACCCGCGCGATCGCACGCTTTGAAAGAGGTGGAGCGACTGAAGAGAAATC 1329
Db 181 GCGCGCTAGTACCCGCGCGATCTCACGCTTTGAAAGAGGTGGAGCGACTGAAGAGAAATC 240
QY 1330 GCGAGAAGCGACGCGCCCGCAGCGCCGAGATGAAGAGAGAGAGAGTGGCGCTGATGAACC 1389
Db 241 GCGAGAAGCGACGCGCCCGCAGCGCCGAGATGAAGAGAGAGAGAGTGGCGCTGATGAACC 300
QY 1390 AGGATCCGCGCAATCCAAACTGCGAGACGCGCGCAATGATACGCGAATATCAGAGCACGC 1449
Db 301 AGGATCCGCGCAATCCAAACTGCGAGACGCGCGCAATGATACGCGAATATCAGAGCACGC 360
QY 1450 TGGAAATTTGTGCGCTGCTCGATGGCGAGCGCGTTCGATGACCATCAGATCACAAGTGTGCG 1509
Db 361 TGGAAATTTGTGCGCTGCTCGATGGCGAGCGCGTTCGATGACCATCAGATCACAAGTGTGCG 420
QY 1510 TGGCGAAGCGTCTCCATTAGCGCAAGAGGTCAATCGCAAGAGAGATCGATGTCAATTCGG 1569
Db 421 TGGCGAAGCGTCTCCATTAGCGCAAGAGGTCAATCGCAAGAGAGATCGATGTCAATTCGG 480
QY 1570 TGC CGCGCAAGGACATGCTCATCGTCAGCGCGCGCAGCAAGGTTCGACCTCACCAAGT 1629
Db 481 TGC CGCGCAAGGACATGCTCATCGTCAGCGCGCGCAGCAAGGTTCGACCTCACCAAGT 540
QY 1630 TCTGTGAGAACCAACCAAGTTTCGCTTCGACTAGCGCTTCAACGACACGTCGCGCAATGCCA 1689
Db 541 TCTGTGAGAACCAACCAAGTTTCGCTTCGACTAGCGCTTCAACGACACGTCGCGCAATGCCA 600
QY 1690 TGTATATAAATACACAGCAAGCGCTTGTGTGAAAACCAATTTTCGAGGGCGGAATGGCGA 1749
Db 601 TGTATATAAATACACAGCAAGCGCTTGTGTGAAAACCAATTTTCGAGGGCGGAATGGCGA 660
QY 1750 CGTGCTTCGCTTACGCCAGACGGGATCGGGCAAAACGACACACCTACCGCGGTGAGTTTA 1809
Db 661 CGTGCTTCGCTTACGCCAGACGGGATCGGGCAAAACGACACACCTACCGCGGTGAGTTTA 720
QY 1810 ATGGAAGGTGCGAGGACTGCAAGAACCGCATCTACGCCATGGC 1852
Db 721 ATGGAAGGTGCGAGGACTGCAAGAACCGCATCTACGCCATGGC 763

RESULT 5
BF489668
LOCUS
DEFINITION
BF489668 722 bp mRNA linear EST 23-APR-2001
AT25778.5prime AT Drosophila melanogaster adult testes pOTB7
Drosophila melanogaster cDNA clone AT25778 5 similar to CG1453:
FBan0001453 'motor protein' located on: X 10A6-10A7.; 04/09/2001,
mRNA sequence.
ACCESSION
BF489668
VERSION
BF489668.2 GI:13755317
KEYWORDS
EST.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 722)
REFERENCE
AUTHORS
Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Baxter,E.,
Berman,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V.,
Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N.,
Li,P., Liao,G., Miranda,A., Misra,S., Mungall,C.J., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C.,
Lewis,S.E., Celnikier,S. and Rubin,G.M.
BDGP/HMI AT Drosophila EST Project
Unpublished (2000)
TITLE
BDGP/HMI AT Drosophila EST Project
JOURNAL
On Dec 6, 2000 this sequence version replaced gi:11572969.
COMMENT
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
```

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST>, [est@fruitfly.berkeley.edu](mailto:est@fruitfly.berkeley.edu)  
hit genomic AE003485: arm:X [10878159,11180508]

estimated-cyto:10A3-10B11: 04/09/2001

Plate: AT.257 row: G column: 6

High quality sequence stop: 720.

Location/Qualifiers

## FEATURES

source

1..722

/organism="Drosophila melanogaster"

/mol\_type="mRNA"

/db\_xref="taxon:7227"

/clone="AT25778"

/sex="male"

/dev\_stage="0-3 day old Ore-R males"

/lab\_host="Plates AT.10-AT.120: DH5-alpha. Plates

AT.121-AT.319: DH5-alpha Tona"

/clone\_lib="AT Drosophila melanogaster adult testes pOTB7"

/note="Organ: ADULT testes; Vector: pOTB7; Site: 1: EcoRI;

Site 2: XhoI; The mRNA for the testis library was made

from testes and seminal vesicles hand dissected from 0-3

day old Ore-R males. RNA kindly provided by the lab of

Margaret Fuller. Sized fractionated cDNAs were directly

ligated into pOTB7. Plasmid cDNA library."

## ORIGIN

Query Match 23.4%; Score 720.4; DB 2; Length 722;

Best Local Similarity 99.9%; Pred. No. 5.9e-176;

Matches 721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 423 AAATATTATTCGACCTAAATAATCTGACTAAATGTCTTTTTCGTTTATGTTATTTATTTA 482
DB 1 AAATATTATTCGACCTAAATAATCTGACTAAATGTCTTTTTCGTTTATGTTATTTATTTA 60
QY 483 GGCACATTTTGCACACCAACGAGTAGTTACTACATCTACGACTAACCGAACTCTCTCTGC 542
DB 61 GGCACATTTTGCACACCAACGAGTAGTTACTACATCTACGACTAACCGAACTCTCTCTGC 120
QY 543 AAGCAGTGGAGTTGCTGTCCATCAAGCAGTACTCGAGTTAAACGAGGATAAGCCGGGA 602
DB 121 AAGCAGTGGAGTTGCTGTCCATCAAGCAGTACTCGAGTTAAACGAGGATAAGCCGGGA 180
QY 603 GAAGAGAAAGAGATCGGTGGAGATAGAGATATACAGGTGGAGTCAAGAGGAGGATC 662
DB 181 GAAGAGAAAGAGATCGGTGGAGATAGAGATATACAGGTGGAGTCAAGAGGAGGATC 240
QY 663 ATGGACATGATTACGTTGGGCGACAGCGTCAAGATCAAGCGGAGCGGATGCCGCTCCAC 722
DB 241 ATGGACATGATTACGTTGGGCGACAGCGTCAAGATCAAGCGGAGCGGATGCCGCTCCAC 300
QY 723 ATGCCCGTGGTGGCGGTGATCAACAGTTCGGGCAAGTGTCATCAGATCGATCGTACGAG 782
DB 301 ATGCCCGTGGTGGCGGTGATCAACAGTTCGGGCAAGTGTCATCAGATCGATCGTACGAG 360
QY 783 CGCGGCGAAACGAGGCGCAAGGAGTGAAGTGAACGACCCATCTACGCTCAATCCGGAG 842
DB 361 CGCGGCGAAACGAGGCGCAAGGAGTGAAGTGAACGACCCATCTACGCTCAATCCGGAG 420
QY 843 CTAATGCAAGATCTGTCGAACAGCAGCGCGCCCGGAGCCCAAGAAACAGCCCGCG 902
DB 421 CTAATGCAAGATCTGTCGAACAGCAGCGCGCCCGGAGCCCAAGAAACAGCTACCGCG 480
QY 903 CCGATGAACCTCTCGCTTAATCCCAACAATCGGCTATCGGTGGCAATCTCACCAGCCGT 962
DB 481 CCGATGAACCTCTCGCTTAATCCCAACAATCGGCTATCGGTGGCAATCTCACCAGCCGT 540
QY 963 ATGACCATGGCCGAAACATGCTGAACAAGATCCAGGAAGCCGATCGATTCGCAATCCG 1022
DB 541 ATGACCATGGCCGAAACATGCTGAACAAGATCCAGGAAGCCGATCGATTCGCAATCCG 600
QY 1023 ATTGTACGACGAATACGCGTGAATACAAACAGCAACTCCAACTACGCGCGCGGAGGT 1082
DB 601 ATTGTACGACGAATACGCGTGAATACAAACAGCAACTCCAACTACGCGCGCGGAGGT 660
```

```
QY 1083 GGTGGCACCAACAGCTCGACGACACCTGGATTACAGGCTCCACGGTACTCGCAAGCTGCT 1142
DB 661 GGTGGCACCAACAGCTCGACGACACCTGGATTACAGGCTCCACGGTACTCGCAAGCTGCT 720
QY 1143 AC 1144
DB 721 AC 722
```

## RESULT 6

DR122830

LOCUS

DEFINITION

49335266 Drosophila pseudoobscura embryonic cDNA library Drosophila

pseudoobscura cDNA clone G9 5', mRNA sequence.

DR122830

VERSION

DR122830.1

KEYWORDS

EST.

SOURCE

Drosophila pseudoobscura

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 1105)

AUTHORS

Richards, S., Liu, Y., Bettencourt, B. R., Hradecky, P., Letovsky, S.,

Nielsen, R., Thornton, K., Hubisz, M. J., Chen, R., Meisel, R. P.,

Couronne, O., Hua, S., Smith, M. A., Zhang, P., Liu, J., Bussmaker, H. J.,

van Batenburg, M. F., Howells, S. L., Scherer, S. E., Sodergren, E.,

Matthews, B. B., Crosby, M. A., Schroeder, A. J., Ortiz-Barrientos, D.,

Rives, C. M., Metzker, M. L., Muzny, D. M., Scott, G., Steffen, D.,

Wheeler, D. A., Worley, K. C., Havlak, P., Durbin, K. J., Egan, A.,

Gill, R., Hume, J., Morgan, M. B., Miner, G., Hamilton, C., Huang, Y.,

Waldron, L., Verdusco, D., Clerc-Blankenburg, K. P., Dubchak, I.,

Noor, M. A., Anderson, W., White, K. P., Clark, A. G., Schaeffer, S. W.,

Gelbart, W., Weinstock, G. M. and Gibbs, R. A.

Comparative genome sequencing of Drosophila pseudoobscura:

chromosomal, gene, and cis-element evolution

Genome Res. 15 (1), 1-18 (2005)

15632085

COMMENT

Contact: Stephen Richards

Human Genome Sequencing Center

Baylor College of Medicine

One Baylor Plaza, Houston, TX 77030, USA

Tel: 713-798-6667

Email: [stephenr@bcm.tmc.edu](mailto:stephenr@bcm.tmc.edu)

NCBI Trace Archive: 226720539

Insert Length: 1750 Std Error: 0.25.

Location/Qualifiers

1..1105

/organism="Drosophila pseudoobscura"

/mol\_type="mRNA"

/db\_xref="taxon:7237"

/clone="G9"

/dev\_stage="0-18h embryos"

/clone\_lib="Drosophila pseudoobscura embryonic cDNA

library"

/notes="Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; oligo

dt priming from poly A+ RNA, directionally cloned"

ORIGIN

Query Match 23.3%; Score 718; DB 8; Length 1105;

Best Local Similarity 83.0%; Pred. No. 2.8e-172;

Matches 853; Conservative 0; Mismatches 175; Indels 3; Gaps 3;

QY 1466 GCTCGATGGCCGCGCGTGCATGACCATCAGATCAGAGTGTGCGTGGCGCAAGCTCCCA 1524

DB 2 GCACGGGGTGCAGCGGTGACGATCACCAGATCACCGCTTTTGTGCGCAAGCTCCAC 61

QY 1525 TTAGCCGCAAGGAGGTCAATCGCAAGGAGATCGATTCATTCGTCGCCGCAAGGACA 1584

DB 62 CAAGCCGAAGGAGGTGATCCCTTAAGAGGTGATGTCATATCGTGGTCCCGCAGGACT 121

QY 1585 TGCTCATCGTGCACGAGCGCGCAGCAAGGTGCGACCTCCACCAAGTTCTCTGGAGAACACA 1644

Db 122 TGATGATTGTCATAGCGCGGACCAAGGTGCGATCTCACCAAGTTCTTGAACATCAC 181  
Qy 1645 AGTTTCGCTTCGACTAGCCCTTCAACGACAGTGCAGCAATGCGATGATACAAATACA 1704  
Db 182 AATTTCGTTTCGACTATGCTTTCAACGAAATGTGTGCAATGCGATGCTGTACAAATACA 241  
Qy 1705 CAGCAAGCGTTGGTGAACACCATTTTCAGAGGCGGAATGGGAGTCTTCGCTCAAG 1764  
Db 242 CAGCAAGCCCTGGTGAACACTATTTTCAGGCGGAATGGGAGTCTTCGCTCAAG 301  
Qy 1765 GCCAGACGGGATCGGGCAACGACACCATGCGGCGGTGAGTTTAATGGAAGGTGCAGG 1824  
Db 302 GCCAGACAGGATCGGCAAGACCATACCATGSGCGGGAGTTTACGGGTAAATACAGG 361  
Qy 1825 ACTGCAAGACGGGATCTAGCCATGCGGCGCAAGATGTCTTTGTGACCTGAATATGC 1884  
Db 362 ACTGCAAGACGGGATCTATGSCCATGCGGCCCAAGGATGTGTTCCGCGACCTCAATAC 421  
Qy 1885 CGGTTTACCGCGCCATGAATCTAGTCTCTCGGCCAGTTTCTTTTGAGATTTACAGTGGCA 1944  
Db 422 CGGCTATCGACCATGAATCTAGTCTCTCGGCCAGTTTCTTTTGAGATTTACAGTGGCA 481  
Qy 1945 AGGTCTTCGATCTTCTGTCGCAAGCAGAAACTCGCGCTCTCGAGGATGTAACAGC 2004  
Db 482 AGGTCTTCGATCTGTCGCGCGATAAGCAGAAAGCTCGCGCTCTCGAGGACGGCAACAGC 541  
Qy 2005 AAGTGCAGGTGGTGGACTACCGAGNAGTGTGATGCGCTCGAGGAGTACTGAGC 2064  
Db 542 AAGTGCAGGTGGTGGCTTCCAGGAAAGTTTGGACAGCGTGGAGGAGTCTAAAGA 601  
Qy 2065 TCATCCAGCAGCGCAATGCTGCCGCAATCTCGGCCAGAGCTCGGCCCACTCCAATTCTGT 2124  
Db 602 TCATCCAGCAGGGAATGCGGCGCACCTCGGGCGAGACTCGGCCAATGCGAATCTCGT 661  
Qy 2125 CGGCTTCGCAAGCGCGTTTTCAGATTTGCTGCGGCGCGAGGCTCGACGAAGATCCATG 2184  
Db 662 CGCGATCGCATGCGCTCTCCAGATTGCTCGCTCGGATGGGACGACAAAGATCCATG 721  
Qy 2185 GCAAGTTCTGTTATGATCTGGCGGCAATGAGGCGGCGTGGACACTTCTCGGCGG 2244  
Db 722 GTAAATTTTCTGTTCAATGATCTGGCGGCAATGAGGCGGCTGGACACTCTCGCGCGG 781  
Qy 2245 ATCGGACAGCGGTATGAGGCTGCGGAGATTAACTGCTGCTGGCCCTCAAGAGT 2304  
Db 782 ATCGGACAGCGGTATGAGGCTGCGGAGATTAACTGCTGCTGGCCCTCAAGAGT 841  
Qy 2305 GCATTCGTGCTTGGGCAACAGTTCGGGCCCACTTGGCCCTTCCGCTGCTCCAAACTCACCC 2364  
Db 842 GCATTCGGGCAATTGGG-TAACAGTTCGGGCCCACTTGGCCCTTCCGCTGCTCCAAACTCACCC 900  
Qy 2365 AGGTGCTGCGGACTGTTTCAATGCGGAGAGAGAGAGTGCATGATGATGATGATGAT 2424  
Db 901 ANGTGCTGCGGACTGTTTCAATGCGGAGAGAGAGAGTGCATGATGATGATGATGAT 960  
Qy 2425 CGCGGAGACTTACTCTCTGAGCAGACGCTCAACGCTGCGCTATGCGGATCGGTGCA 2484  
Db 961 CGCGGTG-CAACATTTCTGTGAGCATACGCTCAACGCTGCGCTATGCGGATCGGGTTC 2548  
Qy 2485 AGGAGCTG 2492  
Db 1020 AGGAACGTG 1027

RESULT 7  
DR135726 1153 bp mRNA linear EST 16-JUN-2005  
LOCUS 49291254 Drosophila pseudoobscura embryonic cDNA library Drosophila  
DEFINITION pseudoobscura cDNA clone 13 5', mRNA sequence.  
ACCESSION DR135726  
VERSION DR135726.1 GI:67880816  
KEYWORDS EST.  
SOURCE Drosophila pseudoobscura

ORGANISM

Drosophila pseudoobscura  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 1153)

AUTHORS

Richards,S., Liu,Y., Bettencourt,B.R., Hradecky,P., Letovsky,S.,  
Nielsen,R., Thornton,K., Hubisz,M.J., Chen,R., Meisel,R.P.,  
Couronne,O., Hua,S., Smith,M.A., Zhang,P., Liu,J., Bussemaker,H.J.,  
van Batenburg,M.F., Howells,S.L., Scherer,S.E., Sodergren,E.,  
Matthews,B.B., Crosby,M.A., Schroeder,A.J., Ortiz-Barrientos,D.,  
Rives,C.M., Metzker,M.L., Muzny,D.M., Scott,G., Steffen,D.,  
Wheeler,D.A., Worley,K.C., Havlak,P., Durbin,K.J., Egan,A.,  
Gill,R., Hume,J., Morgan,M.B., Miner,G., Hamilton,C., Huang,Y.,  
Waldron,L., Verdusco,D., Clerc-Blankenburg,K.P., Dubchak,I.,  
Koor,M.A., Anderson,W., White,K.P., Clark,A.G., Schaeffer,S.W.,  
Gelbart,W., Weinscock,G.M. and Gibbs,R.A.

TITLE

Comparative genome sequencing of Drosophila pseudoobscura:  
chromosomal, gene, and cis-element evolution

JOURNAL

Genome Res. 15 (1), 1-18 (2005)

PUBMED

15632085

COMMENT

Contact: Stephen Richards  
Human Genome Sequencing Center  
Baylor College of Medicine  
One Baylor Plaza, Houston, TX 77030, USA  
Tel: 713-798-6667  
Email: stephenr@bcm.tmc.edu

NCBI Trace Archive: 226715462

Insert Length: 1750 Std Error: 0.25.

FEATURES

source

Location/Qualifiers

1..1153

/organism="Drosophila pseudoobscura"

/mol\_type="mRNA"

/db\_xref="taxon:7237"

/clone="I3"

/dev\_stage="0-18h embryos"

/clone\_lib="Drosophila pseudoobscura embryonic cDNA

library"

/note="Vector: pOTB7; Site 1: EcoRI; Site 2: XhoI; oligo

dt priming from poly A+ RNA, directionally cloned"

ORIGIN

Query Match 22.5%; Score 692; DB 8; Length 1153;  
Best Local Similarity 82.9%; Pred. No. 1.7e-168;  
Matches 813; Conservative 0; Mismatches 166; Indels 2; Gaps 2;

Qy 1452 GAATTTGTCGCGTGTCTGATGCGCAGCGCGTGTGACGATCACGATCACCCCTTTATGT 78

Db 19 GACCTGTCGNATCGGCACGAGGTCGCGGTTGACGATCACGATCACCCCTTTATGT 78

Qy 1511 GCGCAAGCGTCCCATTTAGCGGCAAGGAGTCAATCGCAAGGAGATCGATGTCATTTCCGT 1570

Db 79 GCGCAAGCGTCCCAAGCGGAAAGGAGTGAATCTTTAGAGGTGATGTCATATCGGT 138

Qy 1571 GCGCGCAAGGACATCTCTCATCGTGACGAGCGCGCAGCAGGTCGACCTCACCAAGTT 1630

Db 139 GCGCGCAAGGACTTGATGATTTGTCATGCGCGGACCAAGTCTGATCTCACCAAGTT 198

Qy 1631 CTGTGAGAACCAAGTTTCGTTTCGACTTCAACGACACGTCGTCGACCAATGCCAT 1690

Db 199 CCTTGAACATCAAAATTTCTGTTTCGACTATGCTTCAACGAAATGTTGTGACAATGCCAT 258

Qy 1691 GGTATACAAATACAGAGCGGCGGATTCGGTGAACCAATTTTCAGGCGGGAATGCGGAC 1750

Db 259 GGTGTACAAATACAGAGCGGCGGATTCGGTGAACCAATTTTCAGGCGGGAATGCGGAC 318

Qy 1751 GTGCTTCGCTACCGGCACAGCGGATTCGGGCAAAACGACACCATCGGCGGTGAGTTTAA 1810

Db 319 ATGCTTTGCTATGGCCACAGAGGATCCGCAAGACCCATACCATGCGGCGGAGTTTAC 378

Qy 1811 TGGAAAGGTGCAGGACTGCAAGAACGGCATCTACGCCATGGGCGGCAAGATGCTTTGT 1870

Db 379 GGGTAAATACAGGACTGCAAGAACGGCATCTATGCCATGGCGGCCCAAGATGTTTCG 438

```
QY 1871 GACCCTGAATGCGCGTTACCGCCATGAATCTAGTCTGCTCGGCCAGTCTTTTGA 1930
Db 439 GACCCCTCAATACACCGCGCTATCGCACCATGAATCTAGTCTGCTCGGCCCTGTTTGA 498
QY 1931 GATTTACAGTGGCAAGTCTTCGATCTTCTGTCGACCAAGCAGAACTGCGCGTCTCGA 1990
Db 499 GATTTACAGTGGCAAGTCTTTCGATCTGTCGCGCATGAAGCAGAGCTGCGCGTCTCGA 558
QY 1991 CGATGGTAAACGACAGTGCAGGTGGTGGGATCTACCGAGAAAGTGGTGCATGGCGTGA 2050
Db 559 GGACGGCAACGACGAGGTGCAGGTGGTGGGCTCACCGAAAGGTTGTGGACAGCGTGA 618
QY 2051 GGAGGTACTGAAGCTCATCCAGCAGCGCAATGCTGCCGACATCCCGCCAGAGTGGCG 2110
Db 619 GGAGGTGCTAAAGATCATCCAGCAGCGCAATGCGCGGCGACCTCGGGCCAGACTCGGC 678
QY 2111 CAACCTCCAATTCGTGCGCTTCGACGCGCTTTTCAGATTGTGTCGCGCGCGAGGGCTC 2170
Db 679 CAATGCGAACTCGTCGCGATCGCATCGCTCTCCAGATTGTCTCGGTCCGATGGGCAC 738
QY 2171 GACGAGATCATGCGCAAGTCTGTTTCATCGATCTGGCGGGCAATGAGCGGGCGTGA 2230
Db 739 GACAAAGATCATGTTAAATTTTCGTTCAATGATCTGGCGGGCAATGAGCGGGGTGA 798
QY 2231 CACTTCTCGGCCGATCGCAGACGCGTATGAGGGTCCGAGATTAAACAATCGCTGCT 2290
Db 799 CACTCTGTCGGCCGATCGCCAGACGCGATGGAGGGGGCGGAGATTAAAGTCCCTGCT 858
QY 2291 GGCCCTCAAGGAGTGCATTCGTGCGTTGGGCAACAGTCGCGCCCACTTGCCCTTCGCTG 2350
Db 859 GGCACTCAAGGAGTGCATTCGGGCAATGGG-TAAGTTCGCGCCCTCGCTTCGCTG 917
QY 2351 CTCCAAACTCAACGAGTGTGCGGACTCGTTCAATGGCGAGAGCAGAGTGCAT 2410
Db 918 CTCGAAACTCAACGAGTGTGCGGACTCGTTCAATGGAAAGAGAGCAAGACTTCAT 977
QY 2411 GATAGCCATGATCTCGCGGG 2431
Db 978 TGATACCATGATATCGCGGTG 998

RESULT 8
BF498964
LOCUS
DEFINITION
  BF498964 720 bp mRNA linear EST 19-APR-2001
  Drosophila melanogaster adult testes pOTB7
  Drosophila melanogaster cDNA clone AT13454 5 similar to CG1453:
  FBan0001453 'motor protein' located on: X 10A6-10A7; 04/09/2001,
  mRNA sequence.
ACCESSION
  BF498964
VERSION
  BF498964.2 GI:13690852
KEYWORDS
  EST.
SOURCE
  Drosophila melanogaster (fruit fly)
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
  1 (bases 1 to 720)
  Stapleton, M., Broketein, P., Hong, L., Agbavani, A., Baxter, E.,
  Berman, B., Carlson, J., Champ, M., Chavez, C., Chew, M., Dorsett, V.,
  Farfan, D., Friese, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,
  Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C. J., Nunoo, J.,
  Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
  Lewis, S. B., Ceiniker, S. and Rubin, G. M.
  BDGP/HMI AT Drosophila EST Project
  Unpublished (2000)
  On Dec 6, 2000 this sequence version replaced gi:11582265.
  Contact: Stapleton, M.
  BDGP
  Lawrence Berkeley National Lab
  One Cyclotron Rd, Berkeley, CA 94720, USA
  Fax: 510 486 6798
  Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
  hit genomic AB003485: arm:X [10878159,11180508]
```

```
estimated-cyto:10A3-10B11: 04/09/2001
Plate: AT.134 row: E column: 6
High quality sequence stop: 660.
FEATURES
    source
    1..720
    /organism="Drosophila melanogaster"
    /mol_type="mRNA"
    /db_xref="taxon:7227"
    /clones="AT13454"
    /sex="male"
    /dev_stage="0-3 day old Ore-R males"
    /lab_host="Plates AT.10-AT.120: DHS-alpha. Plates
    AT.121-AT.319: DHS-alpha Tona"
    /clone_lib="AT Drosophila melanogaster adult testes pOTB7"
    /note="Organ: ADULT testes; Vector: pOTB7; Site 1: ECORI;
    Site 2: XhoI; The mRNA for the testis library was made
    from testes and seminal vesicles hand dissected from 0-3
    day old Ore-R males. RNA kindly provided by the lab of
    Margaret Fuller. Sized fractionated cDNAs were directly
    ligated into pOTB7. Plasmid cDNA library."
ORIGIN
    Query Match 22 3% Score 685.8; DB 2; Length 720;
    Best Local Similarity 98.6%; Pred. No. 6e-167;
    Matches 712; Conservative 0; Mismatches 8; Indels 2; Gaps 2;
    QY 387 CCACAGTTTATCGCTCGTCAAGAAACAAACGAAATAAAATTTATTTTCGACCTAAATAATC 446
    Db 1 CCACAGTTTATCGCTCGTCAAGAAACAAACGAAATAAAATTTATTTTCGACCTAAATAATC 60
    QY 447 TGACTAAATTTGTTTTTTTATGTTATTTATTTAGGCACATTTTGCACCAACAACGT 506
    Db 61 TGACTAAATTTGTTTTTTTATGTTATTTATTTAGGCACATTTTGCACCAACAACGT 120
    QY 507 AGTTTCTACATCTACGACTAACCGNACTCTCTGCAAGCAGTGGAGTGTCTGTCATC 566
    Db 121 AGTTTCTACATCTACGACTAACCGNACTCTCTGCAAGCAGTGGAGTGTCTGTCATC 180
    QY 567 AAGCAGTACTCTCGGAGTTTAAACGAGGATAACCGGAGAAAGAGAAAGAGATCGTGGAGA 626
    Db 181 AAGCAGTACTACGAGTTTAAACGAGGATAACCGGAGAAAGAGAAAGATCGTGGAGA 240
    QY 627 ATAGAGATATACAGGTGGAGTCAAGAGGAGGATCATGACATGATTACGTTGGGCGAG 686
    Db 241 ATAGAGATATACAGGTGGAGTCAAGAGGAGGATCATGACATGATTACGTTGGGCGAG 300
    QY 687 AGCGTCAAGATCAAGCGGAGGATGCGCGCTCCACATGCGCGTGTGGCGGTGATCAAC 746
    Db 301 AGCGTCAAGATCAAGCGGAGGATGCGCGCTCCACATGCGCGTGTGGCGGTGATCAAC 360
    QY 747 CAGTCGGGCAAGTGCATCACAGTCAAGTGTGAGCGCGCGGAGAAAGAGGCAAGGAG 806
    Db 361 CAGTCGGGCAAGTGCATCACAGTCAAGTGTGAGCGCGCGGAGAAAGAGGCAAGGAG 420
    QY 807 GTAGAACTGGAAGCCCATCTACGCTCAATTCGGAGCTAATGCAAGATATCTGTCGAACAG 866
    Db 421 GTAGAACTGGAAGCCCATCTACGCTCAATTCGGAGCTAATGCAAGATATCTGTCGAACAG 480
    QY 867 CAGCGCGCGCGGAGCCCAAGAAACAGCAGCGCGCGGATGAACCTCTCGCTGAATCCC 926
    Db 481 CAGCGCGCGCGGAGCCCAAGAAACAGCAGCGCGCGGATGAACCTCTCGCTGAATCCC 540
    QY 927 ACACAACTCGCTATCGTGTGCAATCTCACAGCGGTATGACCATGGCCGGAACATGCTG 986
    Db 541 ACACAACTCGCTATCGTGTGCAATCTCACAGCGGTATGACCATGGCCGGAACATGCTG 600
    QY 987 AACAAAGTCCAGAAAGCCAGTGCATTTCCCAATTCGATTTGTCTCAGAGCAATACGCTGAAT 1046
    Db 601 AACAAAGTCCAGAAAGCCAGTGCATTTCCCAATTCGATTTGTCTCAGAGCAATACGCTGAAT 659
    QY 1047 ACAAACAGCACTCCCAACTACGCGCGCGGAGGTGGTGGCAACCAACAGCTCGAGACC 1106
    Db 660 ACAAACAGCACTCCCAACTACGCGCGCGGAGGTGGTGGCAACCAACAGCTCGAGACC 718
```

```
QY 1107 AC 1108
Db 719 AC 720

BG633810 702 bp mRNA linear EST 23-APR-2001
AT29741.5prime AT Drosophila melanogaster adult testes pOTB7
Drosophila melanogaster cDNA clone AT29741 5 similar to CG1453:
FBan0001453 'motor protein' located on: X 10A6-10A7;: 04/09/2001,
mRNA sequence.
BG633810
BG633810.1 GI:13761347
EST.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 702)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Baxter, E.,
Berman, B., Carlson, J., Champe, M., Chavez, C., Chew, M., Dorsett, V.,
Farfan, D., Frise, E., George, R., Gonzalez, M., Guarin, H., Harris, N.,
Li, P., Liao, G., Miranda, A., Misra, S., Mungall, C. J., Nunoo, J.,
Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K., Yu, C.,
Lewis, S. E., Celniker, S. and Rubin, G. M.
BDGP/HMI At Drosophila EST Project
Unpublished (2000)
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003485: arm:X [10878159,11180508]
estimated-cyto:10A3-10B11: 04/09/2001
Plate: AT 297 row: D column: 5
High quality sequence stop: 668.

FEATURES
source
1. 702
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT29741"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha TonA"
/clone_lib="AT Drosophila melanogaster adult testes pOTB7"
/notes="Organ: ADULT testes; Vector: pOTB7; Site: 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into pOTB7. Plasmid cDNA library."

ORIGIN
Query Match 22.2%; Score 683.6; DB 2; Length 702;
Best Local Similarity 99.3%; Pred. No. 2.2e-166;
Matches 697; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 409 GAAACACGAATAAAATTTATTTTCGACCTAAATAATCTGACTAAATTTGTTTTTGT 468
Db 1 GAAACACGAATAAAATTTATTTTCGACCTAAATAATCTGACTAAATTTGTTTTTGT 60

QY 469 TATGTATTTATTTAGGCAC-ATTTTCACACCAACGAGTAGTTACTACATCTACGACTAA 527
Db 61 TATGTATTTATTTAGGCACNATTTTGCACACCAACGAGTAGTTACTACATCTACGACTAA 120

QY 528 CGGAATCTCTCTCGAAGCAGTGGAAAGTTGCTCTCCATCAAGCAGTACTCGGAGTTAAACG 587
```



```
FEATURES
  source
    High quality sequence stop: 643.
    Location/Qualifiers
      1..687
      /organism="Drosophila melanogaster"
      /mol_type="mRNA"
      /db_xref="taxon:7227"
      /clone="AT04507"
      /sex="male"
      /dev_stage="0-3 day old Ore-R males"
      /lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
      AT.121-AT.319: DH5-alpha Tona"
      /clone_lib="AT Drosophila melanogaster adult testes pOTB7"
      /note="Organ: ADULT testes; Vector: pOTB7; Site_1: EcoRI;
      Site_2: XhoI. The mRNA for the testis library was made
      from testes and seminal vesicles hand dissected from 0-3
      day old Ore-R males. RNA kindly provided by the lab of
      Margaret Fuller. Sized fractionated cDNAs were directly
      ligated into pOTB7. Plasmid cDNA library."
ORIGIN
  Query Match      22.1%; Score 680.6; DB 2; Length 687;
  Best Local Similarity 99.4%; Pred. No. 1.3e-165;
  Matches 683; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
  QY 388 CACAAGTTTATCGCTCGTCAAGAACCAACGAAATATAATTTATTTAGGCACATTTTGACACCAACGTA 507
  Db 1 CACAAGTTTATCGCTCGTCAAGAACCAACGAAATATAATTTATTTAGGCACATTTTGACACCAACGTA 60
  QY 448 GACTAAATTTGTTTTTTTGTATTTATTTATTTAGGCACATTTTGACACCAACGTA 507
  Db 61 GACTAAATTTGTTTTTTTGTATTTATTTATTTAGGCACATTTTGACACCAACGTA 120
  QY 508 GTTACTACATCTACGTAACGGAATCTCTCTCGAAGCAGTGGAAGTTGTCCTCATCA 567
  Db 121 GTTACTACATCTACGTAACGGAATCTCTCTCGAAGCAGTGGAAGTTGTCCTCATCA 180
  QY 568 AGCAGTACTCGAGTTAAGCAGGATTAAGCGGAGAAAGAGAGATCGGTGGAGAA 627
  Db 181 AGCAGTACTCGAGTTAAGCAGGATTAAGCGGAGAAAGAGAGATCGGTGGAGAA 240
  QY 628 TAGAGATATACAGTGGAGTCAAGAGAAAGGATCATGGACATGATTACCGTGGGCGAGA 687
  Db 241 TAGAGATATACAGTGGAGTCAAGAGAAAGGATCATGGACATGATTACCGTGGGCGAGA 300
  QY 688 CGGTCAAGATCAAGCGAGCGATGGCGGTGTCATGGCGGTGATGATCAACC 747
  Db 301 CGGTCAAGATCAAGCGAGCGATGGCGGTGTCATGGCGGTGATGATCAACC 360
  QY 748 AGTCGGCAAGTGCATCACAGTGAATGTTACGAGCGCGGAAACGAGGCAAGGAGG 807
  Db 361 AGTCGGCAAGTGCATCACAGTGAATGTTACGAGCGCGGAAACGAGGCAAGGAGG 420
  QY 808 TAGAATCGAGCGCCATCTACGCTCAATCGGAGCTTAATGCAAGATCTGTGCAACAGC 867
  Db 421 TAGAATCGAGCGCCATCTACGCTCAATCGGAGCTTAATGCAAGATCTGTGCAACAGC 480
  QY 868 ACGCCGCCCGGAGCCCAAGAAACAAGCCACCGCGCCGATGAACCTTCGCGTAATCCCA 927
  Db 481 ACGCCGCCCGGAGCCCAAGAAACAAGCCACCGCGCCGATGAACCTTCGCGTAATCCCA 540
  QY 928 CACAATCGGTATCGGTGGCAATCTACACGCGGTATGACATGGCGGAAACATGCTGA 987
  Db 541 CACAATCGGTATCGGTGGCAATCTACACGCGGTATGACATGGCGGAAACATGCTGA 600
  QY 988 ACAAGATCCAGGAAGCCAGTTCGATTCCTCAATCCGATTGTCAGCAGCAATAGCGTGAATA 1047
  Db 601 ACAAGATCCAGGAAGCCAGTTCGATTCCTCAATCCGATTGTCAGCAGCAATAGCGTGAATA 660
  QY 1048 CAAACAGCAACTCAACACTACGCGCG 1074
  Db 661 CAAACAGCAACTCAACACTACGCGCG 687
```

```
RESULT 11
BF499204
LOCUS
DEFINITION
  BF499204      686 bp      mRNA      linear      EST 19-APR-2001
  AT13721.5Prime AT Drosophila melanogaster adult testes pOTB7
  Drosophila melanogaster cDNA clone AT13721.5 similar to CG1453;
  Fban0001453 'motor protein' located on: X 10A6-10A7;; 04/09/2001,
  mRNA sequence.
ACCESSION
  BF499204
VERSION
  BF499204.2 GI:13691069
KEYWORDS
  EST.
SOURCE
  Drosophila melanogaster (fruit fly)
  Drosophila melanogaster
  Sukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
  1 (bases 1 to 686)
AUTHORS
  Stapleton,M., Brokstein,P., Hong,L., Agbavani,A., Baxter,E.,
  Berman,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V.,
  Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N.,
  Li,P., Liao,G., Miranda,A., Misra,S., Mungall,C.J., Nunoo,J.,
  Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C.,
  Lewis,S.B., Celniker,S. and Rubin,G.M.
  BDGP/HMI AT Drosophila EST Project
  Unpublished (2000)
  On Dec 6, 2000 this sequence version replaced gi:11582505.
  Contact: Stapleton, M.
  BDGP
  Lawrence Berkeley National Lab
  One Cyclotron Rd, Berkeley, CA 94720, USA
  Fax: 510 486 6798
  Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
  hit genomic AEO38485: arm:X [10878159..11180508]
  estimated-cyto:10A3-10B11: 04/09/2001
  Plate: AT.137 row: B column: 9
  High quality sequence stop: 609.
  Location/Qualifiers
    1..686
    /organism="Drosophila melanogaster"
    /mol_type="mRNA"
    /db_xref="taxon:7227"
    /clone="AT13721"
    /sex="male"
    /dev_stage="0-3 day old Ore-R males"
    /lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
    AT.121-AT.319: DH5-alpha Tona"
    /clone_lib="AT Drosophila melanogaster adult testes pOTB7"
    /note="Organ: ADULT testes; Vector: pOTB7; Site_1: EcoRI;
    Site_2: XhoI. The mRNA for the testis library was made
    from testes and seminal vesicles hand dissected from 0-3
    day old Ore-R males. RNA kindly provided by the lab of
    Margaret Fuller. Sized fractionated cDNAs were directly
    ligated into pOTB7. Plasmid cDNA library."
ORIGIN
  Query Match      22.0%; Score 678; DB 2; Length 686;
  Best Local Similarity 99.3%; Pred. No. 6.4e-165;
  Matches 681; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
  QY 388 CACAAGTTTATCGCTCGTCAAGAACCAACGAAATATAATTTATTTAGGCACATTTTGACACCAACGTA 447
  Db 1 CACAAGTTTATCGCTCGTCAAGAACCAACGAAATATAATTTATTTAGGCACATTTTGACACCAACGTA 60
  QY 448 GACTAAATTTGTTTTTTTGTATTTATTTATTTAGGCACATTTTGACACCAACGTA 507
  Db 61 GACTAAATTTGTTTTTTTGTATTTATTTATTTAGGCACATTTTGACACCAACGTA 120
  QY 508 GTTACTACATCTACGTAACGGAATCTCTCTCGAAGCAGTGGAAGTTGTCCTCATCA 567
  Db 121 GTTACTACATCTACGTAACGGAATCTCTCTCGAAGCAGTGGAAGTTGTCCTCATCA 180
  QY 568 AGCAGTACTCGAGTTAAGCAGGATTAAGCGGAGAAAGAGAGATCGGTGGAGAA 627
  Db 181 AGCAGTACTCGAGTTAAGCAGGATTAAGCGGAGAAAGAGAGATCGGTGGAGAA 240
```

/note="vector: pOT2; Site 1: EcoRI; Site 2: XhoI; Sized  
fractionated cDNAs were directly ligated into pOT2.  
Plasmid cDNA library."

241	DB	TAGAGATATACAGTGGAGTCAAGAGGAGGATCATGGACATGATTACGGTGGGGCAGA	300
688	QY	CGGTCAAGATCAACGGACGGATGGCCGCGTCCACATGGCCGTTGGTGGCGGTGATCAACC	747
301	DB	CGGTCAAGATCAACGGACGGATGGCCGCGTCCACATGGCCGTTGGTGGCGGTGATCAACC	360
748	QY	AGTCGGGCAGTGTGATCACAGTCGAATGGTACGAGCGCGGCGGAAACGAAGGGCAAGAGG	807
361	DB	AGTCGGGCAGTGTGATCACAGTCGAATGGTACGAGCGCGGCGGAAACGAAGGGCAAGAGG	420
808	QY	TAGAACTGGAGCCCATCTCACGCTCAATCCGAGGCTAAATCAAGATACTCTTCGAACAGC	867
421	DB	TAGAACTGGAGCCCATCTCACGCTCAATCCGAGGCTAAATCAAGATACTCTTCGAACAGC	480
868	QY	ACGCCGCCCGGAGCCCCAAGAAACAAGCCACCGCGCGATGAACCTCTCGCGTAATCCCA	927
481	DB	ACGCCGCCCGGAGCCCCAAGAAACAAGCCACCGCGCGATGAACCTCTCGCGTAATCCCA	540
928	QY	CACAATCGGCTATCGGTGGCAATCTCACAGCCGATGACCATGGCCGGAACAATGCTGA	987
541	DB	CACAATCGGCTATCGGTGGCAATCTCACAGCCGATGACCATGGCCGGAACAATGCTGA	600
988	QY	ACAAGATCCAGGAAAGCCAGTCGATTCCCAATCCGATTGTCAGCAGCAATAGCGGTGAATA	1047
601	DB	ACAAATCCAGGAAAGCCAGTCGATTCCCAATCCGATTGTCAGCAGCAATAGCTTGAATA	660
1048	QY	CAAAACAGCAACTCCAAACACTACGGCC	1073
661	DB	CAAAACAGCAACTCCAAACACTACGGCC	686

RESULT 12	
Bt635599	
LOCUS	Bt635599 689 bp linear EST 10-SEP-2001
DEFINITION	Spl16989.Sprime SD Drosophila melanogaster Schneider L2 cell culture port2 Drosophila melanogaster cDNA clone Spl16989 5 similar to Kj1p10A: FBAN001453 GO:[kinesin (GO:0005871) ; motor protein (GO:0003774)] located on X 10A6-10A7; 05/18/2001, mRNA sequence.

**KEYWORDS** EST.  
**SOURCE** Br033597.1 G113597869  
**ORGANISM** Drosophila melanogaster (fruit fly)  
Drosophila melanogaster  
Drosophila melanogaster  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Dipera; Brachycera; Muscomorpha;  
Phlebotomidae; Phlebotominae; Phlebotomus; Phlebotomus; Phlebotomus;  
Phlebotomus; Phlebotomus; Phlebotomus; Phlebotomus; Phlebotomus;

**TITLE** BGGP/HMMI Drosophila EST Project  
**JOURNAL** Unpublished (2001)  
**COMMENT** Contact: Stapleton, M.  
BDGP

One Lyncrotron Rd, Berkeley, CA 94720, USA  
Fax: 510 486 6798  
Email: [http://www.fruitfly.org/EST\\_est@fruitfly.berkeley.edu](http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu)  
hit genomic AE003485: arm:X [10878159,1180508]  
estimated-cyto:10A3-10B11: 05/18/2001

```

FEATURES
source
1. .689
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clones="SD16989"
/lab_host="DH5-alpha"
/clone_lib="SP Drosophila melanogaster Schneider L2 cell
culture pot2"

```

ORIGIN	Query Match	21.9%;	Score 674.4;	DB 3;	Length 689;
	Best Local Similarity	99.1%;	Pred. No. 5.6e-164;		
	Matches 678;	Conservative 0;	Mismatches 6;	Indels 0;	Gaps 0;
Qy	385	TTCCACAGTTTATCGCTCGTCAAGAAACAACGAAATAAAATATTTTCGACCTAAAAA	444		
Db	6	TTTCCCAAGTTTATCGCTCGTCAAGAAACAACGAAATAAAATATTTTCGACCTAAAAA	65		
Qy	445	TCCTGACTAAATTGCTGTTTTTTGTATGTATTTATTTAGGCACATTTTGCACACCAAC	504		
Db	66	TCCTGACTAAATTGCTGTTTTTTGTATGTATTTATTTAGGCACATTTTGCACACCAAC	125		
Qy	505	GTAGTTACTACATCTACGACTAACCGAACTCCTCTCGAACGAGTGGAGTTGCTGTCCA	564		
Db	126	GTAGTTACTACATCTACGACTAACCGAACTCCTCTCGAACGAGTGGAGTTGCTGTCCA	185		
Qy	565	TCAGCAGTACTCGGAGTTTAACGACAGTAAAGCCGGGAGAAAGAGAAAGATCGGTGA	624		
Db	186	TCAGCAGTACTACGCGAGTTTAAACGACAGTAAAGCCGGGAGAAAGAGAAATTCGGTGA	245		
Qy	625	GAAATAGAGATATACAGGTGGAGTCAAAAGAGGAAGGATCGGACATGATTTACGGTGGGGC	684		
Db	246	GAAATAGAGATATACAGGTGGAGTCAAAAGAGGAAGGATCGGACATGATTTACGGTGGGGC	305		
Qy	685	AGAGCGTCAAGATCAAGCGGACGGATGGCCGGTCCAATGGCCGTTGGCGGTGATCA	744		
Db	306	AGAGCGTCAAGATCAAGCGGACGGATGGCCGGTCCAATGGCCGTTGGCGGTGATCA	365		
Qy	745	ACCAGTCGGGCAAGTGCATCACAGTTCGAATGGTACGAGCGCGGGGAAAACGAAGGGCAAGG	804		
Db	366	ACCAGTCGGGCAAGTGCATCACAGTTCGAATGGTACGAGCGCGGGGAAAACGAAGGGCAAGG	425		
Qy	805	AGGTAGAACTGGAGCCCATCTACAGCTCAATCCGAGGCTAAATGCAAGATACTGTTCGAAC	864		
Db	426	AGGTAGAACTGGAGCCCATCTACAGCTCAATCCGAGGCTAAATGCAAGATACTGTTCGAAC	485		
Qy	865	AGCACGCGCCCCGGAGCCCCAAGAAAACAAGCCACCGCGCCGATGAACCTCTCGCGTAATC	924		
Db	486	AGCACGCGCCCCGGAGCCCCAAGAAAACAAGCCACCGCGCCGATGAACCTCTCGCGTAATC	545		
Qy	925	CCACACAATCGGCTATCGGTGGCAATCTCACCAGCCGATATGACCATGGCCGGAAAACATGC	984		
Db	546	CCACACAATCGGCTATCGGTGGCAATCTCACCAGCCGATATGACCATGGCCGGAAAACATGC	605		
Qy	985	TGAACAAGATCCAGGAAAACGAGTCGATTCCCAATCCGATTGTTCAGCAGCAATAGCGGTGA	1044		
Db	606	TGAACAAGATCCAGGAAAACGAGTCGATTCCCAATCCGATTGTTCAGCAGCAATAGCGGTGA	665		
Qy	1045	ATACAAACAGCAACTCCAACTA	1068		
Db	666	ATACAAACAGCAACTCCAACTA	689		

RESULT 13	BF495237	679 bp	mRNA	linear	EST 19-APR-2000
LOCUS	AT04114.5				
DEFINITION	Prime AT <i>Drosophila melanogaster</i> adult testes pOTB7 Drosophila melanogaster cDNA, clone AT04114.5 similar to CG1453; FBano001453 "motor protein" located on: X 10A6-10A7;; 04/07/2001, mRNA sequence.				
ACCESSION	BF495237				
VERSION	BF495237.2	GI:13685720			
KEYWORDS	EST.				
SOURCE	<i>Drosophila melanogaster</i> (fruit fly)				
ORGANISM	<i>Drosophila melanogaster</i>				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				

```

REFERENCE
AUTHORS
1 (bases 1 to 679)
Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Baxter,E.,
Berman,B., Carlson,J., Champe,M., Chavez,C., Chew,M., Dorsett,V.,
Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N.,
Li,P., Liao,G., Miranda,A., Misra,S., Mungall,C.J., Nunoo,J.,
Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C.,
Lewis,S.E., Celniker,S. and Rubin,G.M.
BDGP/HMI AT Drosophila EST Project
UNPUBLISHED (2000)
On Dec 6, 2000 this sequence version replaced gi:11578538.
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
hit genomic AE003485: arm:X [10878159,11180508]
estimated-cyto:10A3-10B11: 04/07/2001
Plate: AT.41 row: B column: 2
High quality sequence stop: 655.
FEATURES
source
1. .679
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="AT04114"
/sex="male"
/dev_stage="0-3 day old Ore-R males"
/lab_host="Plates AT.10-AT.120: DH5-alpha. Plates
AT.121-AT.319: DH5-alpha Tona"
/clone_lib="AT Drosophila melanogaster adult testes potB7"
/notes="Organ: ADULT testes; Vector: potB7; Site 1: EcoRI;
Site 2: XhoI; The mRNA for the testis library was made
from testes and seminal vesicles hand dissected from 0-3
day old Ore-R males. RNA kindly provided by the lab of
Margaret Fuller. Sized fractionated cDNAs were directly
ligated into potB7. Plasmid cDNA library."
ORIGIN
Query Match 21.5%; Score 663.6; DB 2; Length 679;
Best Local Similarity 99.0%; Pred. No. 3.2e-161;
Matches 668; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 388 CACAGTTTATCGTCTGCAAGAAACCAACGAAATAAAATTTATTTTCGACCTAAATAATCT 447
DB 1 CACAGTTTATCGTCTGCAAGAAACCAACGAAATAAAATTTATTTTCGACCTAAATAATCT 60
QY 448 GACTAATATGTTTGTGTTATGTTATTTATTTAGGCACATTTTGACACACCAACGTA 507
DB 61 GACTAATATGTTTGTGTTATGTTATTTATTTAGGCACATTTTGACACACCAACGTA 120
QY 508 GTTACTACATCTACGACTAACGGAACTCCTCCTGCAAGCAGTGAAGTGTCTGTCCATCA 567
DB 121 GTTACTACATCTACGACTAACGGAACTCCTCCTGCAAGCAGTGAAGTGTCTGTCCATCA 180
QY 568 AGCAGTACTCGGAGTTAACCGAGATTAAGCCGGAGAAAGAGAGATCGGTGGAGAA 627
DB 181 AGCAGTACACGGAGTTAACCGAGATTAAGCCGGAGAAAGAGAGATCGGTGGAGAA 240
QY 628 TAGAGATATACAGTGTGAGTCAAGAGGAGAGATCATGGACATGATTACCGTGGGCAGA 687
DB 241 TAGAGATATACAGTGTGAGTCAAGAGGAGAGATCATGGACATGATTACCGTGGGCAGA 300
QY 688 CGGTCAAGATCAAGCGGACGGATGGCCCGCTGCACATGGCCGCTGGTGGCGGTGATCAACC 747
DB 301 CGGTCAAGATCAAGCGGACGGATGGCCCGCTGCACATGGCCGCTGGTGGCGGTGATCAACC 360
QY 748 AGTCGGCAAGTGCATCACAGTGCAGTGCAGTGCAGCGCGCGGAAACGAAAGGCAAGGAGG 807
DB 361 AGTCGGCAAGTGCATCACAGTGCAGTGCAGTGCAGCGCGCGGAAACGAAAGGCAAGGAGG 420
QY 808 TAGAATCGGACGCCATCTACGCTCAATCGGAGCTAATGCAAGATCTGTGCAACAGC 867

```

Query Match

21.5%; Score 663.6; DB 2; Length 670;

Best Local Similarity 99.4%; Pred. No. 3.6e-161; Matches 666; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 379 CTTAGTTTCACAAAGTTTATCGCTCGTCAAGAAACAAACGAAATAAAATTTATTTTCACCT 438

Db 1 CTTAGTTTCACAAAGTTTATCGCTCGTCAAGAAACAAACGAAATAAAATTTATTTTCACCT 60

QY 439 AAAAAATCTGACTAAATTTGTTTTTTTATGTATTTATTTAGGCACATTTTGCACAC 498

Db 61 AAAAAATCTGACTAAATTTGTTTTTTTATGTATTTATTTAGGCACATTTTGCACAC 120

QY 499 CACAACGTAGTTACTACATCTACGACTTAACGGAACTCTCTCTCAAGACGATGGAAATTGC 558

Db 121 CACAACGTAGTTACTACATCTACGACTTAACGGAACTCTCTCTCAAGACGATGGAAATTGC 180

QY 559 TGTCCTCATCAAGCAGTACTCGGAGTTAAACGAGATAAGCCGGGAGAAAGAAAGATC 618

Db 181 TGTCCTCATCAAGCAGTACTCGGAGTTAAACGAGATAAGCCGGGAGAAAGAAAGATC 240

QY 619 GGTGGAGATAGAGATATACAGGTGGAGTCAAGAGAGAAAGGATCATGGACATGATTACGS 678

Db 241 GGTGGAGATAGAGATATACAGGTGGAGTCAAGAGAGAAAGGATCATGGACATGATTACGS 300

QY 679 TGGGGCAGAGCGTCAAGATCAACGGACGAGTGGCGCGTCCACATGGCGGTGGTGGCGG 738

Db 301 TGGGGCAGAGCGTCAAGATCAACGGACGAGTGGCGCGTCCACATGGCGGTGGTGGCGG 360

QY 739 TGATCAACCAAGTCGGGCAAGTGATCACAGTCAAGTGGTACGAGCGCGGCGAAACGAAAG 798

Db 361 TGATCAACCAAGTCGGGCAAGTGATCACAGTCAAGTGGTACGAGCGCGGCGAAACGAAAG 420

QY 799 GCAAGGAGGTAGAACTGGAGCGCATCTACGCTCAATCCGAGCTAATGCAAGATACG 858

Db 421 GCAAGGAGGTAGAACTGGAGCGCATCTACGCTCAATCCGAGCTAATGCAAGATACG 480

QY 859 TCGAACAGCAGCGCGCGCGCGGAGCCCAAGAAACAAGCCACCGCGCGGATGAACCTCTCG 918

Db 481 TCGAACAGCAGCGCGCGCGCGGAGCCCAAGAAACAAGCCACCGCGCGGATGAACCTCTCG 540

QY 919 GTAATCCACACAAATCGGCTATCGGTGGCATCTACAGCGCGTATGACATGGCGGAA 978

Db 541 GTAATCCACACAAATCGGCTATCGGTGGCAATCTCACAGCGGTATGACCATGGCGGAA 600

QY 979 ACATGCTGAACAGATCCAGGAAAGCCAGTCGATTCCCAATCCGATTGTTCAGCAGCAATA 1038

Db 601 ACATGCTGAACAAATCCAGAAAGCCAGTCGATTCCCAATCCGATTGTTCAGCAGCAATA 660

QY 1039 GCGTGAATAC 1048

Db 661 GCTTGAATAC 670

RESULT 15

BI609735

LOCUS

DEFINITION

BI609735 664 bp mRNA linear EST 07-SEP-2001

RH14556.5prime RH Drosophila melanogaster normalized Head pFlc-1

Drosophila melanogaster cDNA clone RH14556 5 similar to Klp10A:

PB0001453 GO:[kinesin (GO:0005871); motor protein (GO:0003774)]

located on: X 10A6-10A7; 08/17/2001, mRNA sequence.

BI609735

BI609735.1 GI:15505260

EST.

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 664)

Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Mira,S., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and

Rubin,G.M.

BDGP/RHMI RH Drosophila EST Project

Unpublished (2001)

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu

hit genomic AB003485: arm:X [10878159,11180508]

estimated-cyto:10A3-10B11: 08/17/2001

Plate: RH.145 row: E column: 8

High quality sequence stop: 563.

Location/Qualifiers

1..664

/organism="Drosophila melanogaster"

/mol\_type="mRNA"

/db\_xref="taxon:7227"

/clone="RH14556"

/sex="male and female"

/dev\_stage="Adult"

/lab\_host="DHS-alpha TonA"

/clone\_lib="RH Drosophila melanogaster normalized Head pFlc-1"

/note="Organ: head; Vector: pFlc1; Site 1: XhoI; Site 2: BamHI; Library was kindly generated by Piero Carninci at the RIKEN. The library was normalized and excised using Cre recombinase. Plasmid cDNA library."

ORIGIN

Query Match 21.5%; Score 662.4; DB 3; Length 664;

Best Local Similarity 99.8%; Pred. No. 7.4e-161;

Matches 663; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 506 TAGTTACTACATCTACGACTAAACGAACTCTCTCTCAAGACGATGGAGTTGCTGTCCAT 565

Db 1 TAGTTACTACATCTACGACTAAACGAACTCTCTCTCAAGACGATGGAGTTGCTGTCCAT 60

QY 566 CAAGCAGTACTCGGAGTTAAACGAGATAAGCCGGGAGAAAGAGATCGGTGGAG 625

Db 61 CAAGCAGTACTCGGAGTTAAACGAGATAAGCCGGGAGAAAGAGATCGGTGGAG 120

QY 626 AATAGAGATATACAGTGGAGTCAAGAGAGAGAGATCATGGACATGATTACGGTGGGCA 685

Db 121 AATAGAGATATACAGTGGAGTCAAGAGAGAGAGATCATGGACATGATTACGGTGGGCA 180

QY 686 GAGCCTCAAGATCAAGCGGACGAGATGGCGCGTCCACATGGCCGTGGTGGCGGTGATCAA 745

Db 181 GAGCCTCAAGATCAAGCGGACGAGATGGCGCGTCCACATGGCCGTGGTGGCGGTGATCAA 240

QY 746 CCAGTCGGGCGAGTGCATCAGTCGAAATGGTACGAGCGCGGCGGAAACGAGGCGAAGA 805

Db 241 CCAGTCGGGCGAGTGCATCAGTCGAAATGGTACGAGCGCGGCGGAAACGAGGCGAAGA 300

QY 806 GGTAGAACTGGAAGCGCATCTCACCTCAATCCGAGCTAATGCAAGATACTGTCCGAACA 865

Db 301 GGTAGAACTGGAAGCGCATCTCACCTCAATCCGAGCTAATGCAAGATACTGTCCGAACA 360

QY 866 GCACGCCGCCCGGAGCCCAAGAAACAAGCCACCGCGCGATGAAACCTCTCGCGTAATCC 925

Db 361 GCACGCCGCCCGGAGCCCAAGAAACAAGCCACCGCGCGATGAAACCTCTCGCGTAATCC 420

QY 926 CACACAATCGGTATCGGTGGCAATCTCACAGCGGTATGACCATGGCGGAAACATGCT 985

Db 421 CACACAATCGGTATCGGTGGCAATCTCACAGCGGTATGACCATGGCGGAAACATGCT 480

QY 986 GAAACAAGATCCAGGAAGCCAGTCCGATTCCCAATCCGATTGTTCAGCAGCAATAGCGTAA 1045

Db 481 GAAACAAGATCCAGGAAGCCAGTCCGATTCCCAATCCGATTGTTCAGCAGCAATAGCGTAA 540

QY 1046 TACAACAGCAACTCTCAACACTACCGCGCGGAGGTGGTGGCAACACCACTTCGACGAC 1105

Db 541 TACAACAGCAACTCTCAACACTACCGCGCGGAGGTGGTGGCAACACCACTTCGACGAC 600

Qy	1106	CACTGGATTACAGCGTCCACGGTACTCGCAAGCTGCTACCGGCCAGCAGCAGACAAGCAT	1165
Db	601	CACTGGATTACAGCGTCCACGGTACTCGCAAGCTGCTACCGGCCAGCAGCAGACAAGCAT	660
Qy	1166	CGCC	1169
Db	661	CGCC	664

Search completed: May 12, 2006, 21:09:21  
 Job time : 11794 secs

***This Page Blank (uspto)***